

# THEORETICA CHIMICA ACTA

edenda curat

**Hermann Hartmann**

adiuvantibus

**C. J. Ballhausen, København**

**R. D. Brown, Clayton**

**E. Heilbronner, Basel**

**J. Jortner, Tel-Aviv**

**M. Kotani, Tokyo**

**J. Koutecký, Praha**

**J. W. Linnett, Cambridge**

**E. E. Nikitin, Moskva**

**R. G. Pearson, Evanston**

**B. Pullman, Paris**

**K. Ruedenberg, Ames**

**C. Sandorfy, Montreal**

**M. Simonetta, Milano**

**O. Sinanoğlu, New Haven**



**Vol. 31, 1973**

**Springer-Verlag · Berlin · Heidelberg · New York**

The exclusive copyright for all languages and countries, including the right for photomechanical and any other reproductions, also in microform, is transferred to the publisher.

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Alle Rechte, einschließlich das der Übersetzung in fremde Sprachen und das der fotomechanischen Wiedergabe oder einer sonstigen Vervielfältigung, auch in Mikroform, vorbehalten. Jedoch wird gewerblichen Unternehmen für den innerbetrieblichen Gebrauch nach Maßgabe des zwischen dem Börsenverein des Deutschen Buchhandels e. V. und dem Bundesverband der Deutschen Industrie abgeschlossenen Rahmenabkommens die Anfertigung einer fotomechanischen Vervielfältigung gestattet. Wenn für diese Zeitschrift kein Pauschalabkommen mit dem Verlag vereinbart worden ist, ist eine Wertmarke im Betrage von DM 0,40 pro Seite zu verwenden. Der Verlag läßt diese Beträge den Autorenverbänden zufließen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in dieser Zeitschrift berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei betrachtet wären und daher von jedermann benutzt werden dürften.

Springer-Verlag Berlin Heidelberg New York  
Printed in Germany by Brühlische Universitätsdruckerei, Gießen  
© by Springer-Verlag Berlin Heidelberg 1973

389.67

S. 32

389.67

Date 24.4.82

hja

# Index

## Commentationes et Relationes

Adams, D. B., Clark, D. T.: <i>Ab initio</i> Calculations of Core Electron Binding Energies and Shifts in Halomethanes . . . . .	171
Arnaud, R., Faramond-Baud, D., Gelus, M.: Etude théorique de la conformation des complexes moléculaires de type donneur accepteur d'électrons. . . . .	335
Boys, S. F., s. Handy, N. C. . . . .	195
Buck, H. M., s. Dobbelaere, J. R. de, <i>et al.</i> . . . .	95
Cederbaum, L. S.: Direct Calculation of Ionization Potentials of Closed-Shell Atoms and Molecules . . . . .	239
Chaillet, M., s. Liotard, D., <i>et al.</i> . . . .	325
Chen, S.-Y., Hedges, R. M.: Calculations of Singlet and Triplet States of Some Azabenzenes by Modified INDO-CI . . . . .	275
Christoffersen, R. E., s. Shipman, L. L. . . . .	75
Clark, D. T., s. Adams, D. B. . . . .	171
Courrière, P., s. Pullman, B. . . . .	19
Cremašchi, P., Gamba, A., Simonetta, M.: The Influence of Solvation on $\pi^* \leftarrow n$ and $\pi^* \leftarrow \pi$ Transition Energies in Molecules Containing Aza or Carbonyl Groups . . . .	155
Csizmadia, I. G., s. Haines, W. J. . . . .	283
Csizmadia, I. G., s. Hopkinson, A. C. . . . .	83
Csizmadia, I. G., s. Wolfe, S. . . . .	355
Dargelos, A., s. Liotard, D., <i>et al.</i> . . . .	325
Dierksen, G. H. F., Niessen, W. von, Kraemer, W. P.: SCF LCGO MO Studies on the Fluoronium Ion $FH_2^+$ and Its Hydrogen Bonding Interaction with Hydrogen Fluoride $FH$ . . . . .	205
Dobbelaere, J. R. de, Haan, J. W. de, Buck, H. M., Visser, G. J.: CNDO/2 and INDO Calculations of a Reaction Pathway for the Sigmatropic [1,5]H-Shift in Cyclopentadiene . .	95
Faramond-Baud, D., s. Arnaud, R., <i>et al.</i> . . . .	335
Favini, G., Nava, A.: Modes of Interconversion in the Cycloheptene Ring . . . . .	261
Fischer-Hjalmar, I., Siegbahn, P.: A Comparative <i>ab initio</i> Study of Ethylene, Acetylene and Benzene . . . . .	1
Flurry, R. L., Jr.: The Use of Site Symmetry in Constructing Symmetry Adapted Functions . . . . .	221
Gamba, A., s. Cremašchi, P., <i>et al.</i> . . . .	155
Gelus, M., s. Arnaud, R., <i>et al.</i> . . . .	335
Goodisman, J.: Oscillations in the Thomas-Fermi-Dirac Electron Density . . . . .	101
Haan, J. W. de, s. Dobbelaere, J. R. de, <i>et al.</i> . . . .	95
Haines, W. J., Csizmadia, I. G.: An <i>ab initio</i> Molecular Orbital Study on the Addition Reaction of Triplet Nitrene to Ethylene . . . . .	283
Hall, G. G., s. Tait, A. D. . . . .	311
Handy, N. C.: A Method for Selection and Use of Numerical Integration Points for Molecules with Cylindrical Symmetry . . . . .	201
Handy, N. C., Boys, S. F.: Integration Points for the Reduction of Boundary Conditions. .	195

Hase, H. L., Schweig, A.: CNDO/2 (Complete Neglect of Differential Overlap)-Method for Third-Row Molecules . . . . .	215
Hedges, R. M., s. Chen, S.-Y. . . . .	275
Henderson, G. A., Parr, R. G.: Three-Dimensional Bond-Charge Models for Potential Curves of Diatomic Molecules . . . . .	103
Hopkinson, A. C., Csizmadia, I. G.: An <i>ab initio</i> Study of the $A_{Ac}$ I Hydrolysis Mechanism of Formamide . . . . .	83
Hussonnois, M., s. Malý, J. . . . .	137
Hyslop, J.: A Six-Dimensional Quadrature Procedure . . . . .	189
Jug, K.: A New Definition of Atomic Charges in Molecules . . . . .	63
Kitaura, K., Nishimoto, K.: An Application of RPA Theory to Conjugated Systems in the Excited States. II. Linear Polyenes . . . . .	91
Kraemer, W. P., s. Dierksen, G. H. F., <i>et al.</i> . . . .	205
Liotard, D., Dargelos, A., Chaillet, M.: Etude théorique de l'isomérisation <i>syn-anti</i> dans la Formaldoxime . . . . .	325
Lischka, H.: <i>Ab initio</i> Calculations on Small Hydrides Including Electron Correlation IX. Equilibrium Geometries and Harmonic Force Constants of HF, OH <sup>-</sup> , H <sub>2</sub> F <sup>+</sup> and H <sub>2</sub> O and Proton Affinities of F <sup>-</sup> , OH <sup>-</sup> , HF and H <sub>2</sub> O . . . . .	39
Malý, J., Hussonnois, M.: SCF Dirac-Hartree-Fock Calculations in the Periodic System II. Binding Energies and First Ionization Potentials for <i>s</i> , <i>p</i> and <i>d</i> Elements from $Z = 1$ to $Z = 120$ . . . . .	137
Moore, D. J., s. Rees, D., <i>et al.</i> . . . .	183
Nava, A., s. Favini, G. . . . .	281
Nishimoto, K., s. Kitaura, K.: . . . . .	91
Niessen, W. von: A Theory of Molecules in Molecules. II. The Theory and Its Application to the Molecules Be-Be, Li <sub>2</sub> -Li <sub>2</sub> , and to the Internal Rotation in C <sub>2</sub> H <sub>6</sub> . . . . .	111
Niessen, W. von: A Theory of Molecules in Molecules. III. Application to the Hydrogen Bonding Interaction of Two FH Molecules . . . . .	297
Niessen, W. von, s. Dierksen, G. H. F., <i>et al.</i> . . . .	205
Parr, R. G., s. Henderson, G. A. . . . .	103
Port, G. N. J., Pullman, A.: An <i>ab initio</i> Study of the Hydration of Alkylammonium Groups . . . . .	231
Pullman, A., s. Port, G. N. J. . . . .	231
Pullman, B., Courrière, P.: Complementary Molecular Orbital Investigations on the Conformation of Choline Derivatives . . . . .	19
Rees, D., Moore, D. J., Taylor, P. R.: The Numerical Integration of a Molecular Integral Using Two Different Techniques . . . . .	183
Schweig, A., s. Hase, H. L. . . . .	215
Shipman, L. L., Christoffersen, R. E.: <i>Ab initio</i> Calculations on Large Molecules Using Molecular Fragments. Characterization of the Zwitterion of Glycine . . . . .	75
Siegbahn, P., s. Fischer-Hjalmars, I. . . . .	1
Simonetta, M., s. Cremaschi, P., <i>et al.</i> . . . .	155
Stammli, V.: <i>Ab initio</i> Calculations on Small Hydrides Including Electron Correlation X. Triplet-Singlet Energy Separation and Other Properties of the CH <sub>3</sub> Radical . . . . .	49
Tait, A. D., Hall, G. G.: Point Charge Models for LiH, CH <sub>4</sub> , and H <sub>2</sub> O . . . . .	311
Taylor, P. R., s. Rees, D., <i>et al.</i> . . . .	183
Tel, L. M., s. Wolfe, S., <i>et al.</i> . . . .	355
Thomson, C., Wishart, B. J.: Electronic Structure of Unstable Intermediates. III. The Electronic Structure of OCC . . . . .	347



Visser, G. J., s. Dobbelaere, J. R. de, <i>et al.</i> . . . . .	95
Voigt, B.: On Bridging the Gap between the INDO and the NDDO Schemes . . . . .	289
Wagnière, G.: On the Electronic Structure of $N_2H_2$ . A Possible Triplet Ground State in Diazines. . . . .	289
Wishart, B. J., s. Thomson, C. . . . .	347
Wolfe, S., Tel, L. M., Caizmadia, I. G.: Gas Phase Acidities of CH Bonds Adjacent to Oxygen and to Sulphur . . . . .	355

Indexed in Current Contents

## Subjects

<i>Ab initio</i> calculation	1	Electron correlation	39
<i>Ab initio</i> molecular fragment approach	75	Electronic structure	49, 239
Acetylcholine, conformation of	19	Energy analysis	297
—, structure-activity relation of	19	Equivalent cores	171
Addition reaction of triplet nitrene	283	ESCA shifts	171
All-valence electron calculations	215	Excitation energy	49
Approximate LCAO-SCF	289		
Aromaticity of odd-membered cycloradicals	95	Factor groups	221
Atomic and molecular structure	239	Fermi-gas	103
Atomic charges	63	FH...FH hydrogen bonding interaction	297
Azabenzenes	275	$FH_2^+$	205
Aziridine	283	$FH_2^+ \cdot FH$	205
		Force constants	39
Basis set	1	Formaldoxime	325
Be <sub>2</sub>	111		
Benzoecycloheptene, inversion of	261	Gas phase acidity	355
Bond-charge	103	Gaussian quadrature	189
		Geometry of formyl cation	83
Calculated geometry	347	Green's functions	189
Carbanions	355		
C <sub>2</sub> H <sub>2</sub>	111	Hydration of alkylammonium	231
Cheletropic cleavage	269	Hydrides	39
CNDO/2-method for third-row atoms	215	Hydrocarbons	1
Computed activation energy	83	Hydrogen bonding	205
Conformational energy surface	75		
Conjugated systems	91	Ionization	239
Core holes	171	— potentials	1, 275
Cycloheptene conformation	261	— — of atoms	137
		INDO	289
Density oscillations	101		
Diazine	269	Koopmans' theorem	171
Dipole moment formula			
for third-row molecules	215	Li <sub>2</sub>	111
Dipole moments	63	Local anesthetics	19
Donor-acceptor complexes,		Localized orbitals, transferability of	111
conformation of	335		
DRINDO	289	Many-body perturbation theory	239
		Methylene	49
		Molecules in molecules	111, 297
		Multipole-multipole interactions	289

## VI

**NDDO** 289  
**Numerical integration** 183, 195, 201  
**One-electron properties** 311  
**Orbital mapping** 1  
**Oscillator strength** 275  
**Oxime nitrogen inversion** 325  
  
**Pair correlation energies** 49  
**PCILO method** 19  
**Point charge models** 311  
**Polynes** 91  
**Potential curves** 103  
**Projection operators** 221  
**Proton affinity** 39, 355  
**Proton solvation** 205  
**Protonation of Formamide** 83  
  
**Reaction profile** 83  
**RPA calculation** 91

## Index

**SCF LOGO MO calculation** 205  
**Singlet-triplet separation** 269  
**Singularities** 189  
**Site symmetry** 221  
**Solvent shifts of electronic transitions** 155  
**Spin-orbit coupling** 275  
**Symmetry adapted functions** 221  
  
**Thermal sigmatropic hydrogen shifts** 95  
**Thomas-Fermi-Dirac model** 101  
**Transcorrelated method** 201  
**Transferability of localized orbitals** 297  
**Transition energies** 91  
**Transition state geometry** 95  
**Triplet-group state** 269  
  
**Unstable intermediates** 347  
  
**Zwitterion of glycine** 75

## Commentationes

# A Comparative *ab initio* Study of Ethylene, Acetylene and Benzene

Inga Fischer-Hjalmars and Per Siegbahn

Institute of Theoretical Physics, University of Stockholm, Stockholm, Sweden

Received April 3, 1973

Ground state properties have been calculated by use of a medium-sized Gaussian basis set and comparison with other bases has been made. Contraction to "double-zeta" of a comparatively small basis is found to be superior to a large set of primitive Gaussians contracted to minimal basis. Molecular optimization is not important for double-zeta bases. Inclusion of a balanced set of polarization functions is essential in all cases studied. Population analysis gives a certain insight in molecular properties but contour maps are found to be significantly superior. This is demonstrated on bonding properties of corresponding orbitals within the series. In case of benzene Slater's energy-band plot is shown to be useful for classifying bonding properties.

**Key words:** *Ab initio* calculation   Basis set - Hydrocarbons - Ionization potentials - Orbital mapping.

### 1. Introduction

The present paper is concerned with theoretical investigations of ground state properties of small unsaturated organic molecules by means of non-empirical, all-electron calculations. The main goal was to compare calculated properties of the series acetylene, ethylene, benzene, obtained under similar conditions. Several basis sets were used and the effect of various modifications of the basis was studied. The most extensive studies were performed on ethylene. Comprehensive calculations on acetylene were also made, whereas for obvious reasons, applications to benzene were much more restricted.

Section 2 describes the method. In Section 3 the results for ethylene are presented and discussed. Sections 4 and 5 are concerned with acetylene and benzene, respectively. Corresponding orbitals of the different molecules are discussed in Section 6, and Section 7 collects the main results.

### 2. Method

Basis sets of Gaussian type functions localized at the various atoms of the molecule were used. The main part of the study is concerned with calculations employing a basis of nine *s*-type and five *p*-type functions at each carbon atom, (C/9,5), contracted to "double zeta", i.e. two basis functions for each atomic shell, <C/4,2>. For hydrogen, four *s*-type functions, (H/4), were contracted to <H/2>. The effect of including polarization functions, one *d*-function on each carbon

and one  $p$ -function on each hydrogen, was also investigated. Arguments for this type of basis set have been given previously by Roos and Siegbahn [1, 2].

Within this frame an optimal basis set for unsaturated hydrocarbons was determined using ethylene as a probe. A similar investigation has recently been published by Schulman *et al.* [3]. However their study differs from ours in several respects. Most important, their basis set is a minimal one without polarization functions collated to ours of the double-zeta type and including polarization. Furthermore the details of the optimizing procedure were different.

Table 1a. Orbital exponents and contraction coefficients of the Gaussian basis sets. When ethylene-optimized values ( $B'$ ) deviate from atom-optimized,  $B$ , both values are given

Exponent (C/11, 6) [4]	Coefficient	Exponent $B(B')$ : (C/9, 5)	Coefficient
<i>s-type</i>			
469.4	0.000242	—	—
316.47	0.001879	5182.95	0.000947
527.10	0.009743	778.756	0.007256
149.438	0.039167	178.073	0.036292
48.8562	0.123636	50.8779	0.130515
17.6209	0.288316	16.7876	0.320715
6.81082	0.421504	6.14362	0.439516
2.72760	0.254118	2.40398	0.212694
0.756743	1.0	$B$ : 0.511903	1.0
0.300728	1.0	$(B'$ : 0.462000)	1.0)
0.114093	1.0	0.156594	1.0
<i>p-type</i>			
32.6921	0.005869	18.8418	0.013271
7.47261	0.039121	4.15924	0.092479
2.23924	0.151011	$B$ : 1.206710	0.293686
0.772812	0.351470	$(B'$ : 1.136000)	0.293686)
0.274838	1.0	$B$ : 0.385541	0.472915
		$(B'$ : 0.395000)	0.472915)
<i>p<math>\sigma</math></i>			
0.095838	1.0	0.121939	1.0
<i>p<math>\pi</math></i>			
0.095838	1.0	$B$ : 0.121939	1.0
		$(B'$ : 0.131600)	1.0)
<i>d<math>\sigma\sigma</math> = d<math>\sigma\sigma'</math> = d<math>\pi\pi</math></i>			
0.898	1.0	$B$ : 1.000	1.0
		$(B'$ : 0.898)	1.0)
<i>d<math>\sigma\pi</math></i>			
0.700	1.0	$B$ : 1.000	1.0
		$(B'$ : 0.700)	1.0)

Table 1b. See Table 1a for explanation

Exponent	Coefficient	Exponent	Coefficient
A: (H/5, 1) [4]		B': (H/4, 1)	
s-type			
33.865014	0.006068	16.7019	0.019060
5.094788	0.045316	2.51663	0.134240
1.158786	0.202846	0.567196	0.474490
0.325840	1.0	0.154146	1.0
0.102741	1.0		
pσ			
1.200	1.0	1.200	1.0
pπ			
1.000	1.0	1.000	1.0

As starting-point we have chosen the best-atom basis (C/9,5) determined by van Duynenvelt [4] and the (H/4) basis of Huzinaga [5], retaining the linear coefficients determined for the atoms. The contraction used for carbon *s*-type functions was (5, 2, 1, 1) and for *p*-type functions (4, 1). Calculations showed the *p*-type contraction (3, 2) to be considerably inferior. For the *s*-type functions of hydrogen the contraction (3, 1) was used.

During optimization the innermost carbon functions were left unchanged. The exponents of the outer functions, two of *s*-type and three of *p*-type, were varied independently. The outermost *p*-function was chosen differently for  $\sigma$ -orbitals and for  $\pi$ -orbitals. In the case of hydrogen *s*-type functions a scale factor, common to all the exponents, was varied. The polarization functions of hydrogen, *pσ* and *pπ*, were scaled independently. In case of the *d*-type polarization functions of carbon the exponent of the  $d\sigma\pi$ -functions (e.g.  $dxz$  of  $C_2H_4$ , where  $x$  is perpendicular to the plane of the molecule) was varied separate from the other *d*-functions. The final exponents and contraction coefficients are listed in Table 1. It was found that the total energy varied rather slowly with the orbital exponents of the polarization functions. Therefore these exponents could not be determined with high accuracy. The energy was insensitive to small variations above the value 1.25 of the hydrogen *s*-function scale factor. In accord with previous work [2, 6] we have therefore retained the value 1.25 of this factor.

It is interesting to compare the result of the present optimization with that of Schulman *et al.* [3]. These authors found that an ethylene-optimized basis was obtained by contraction of all the best-atom 2s Gaussians by a common scale factor equal to 1.206. The  $2p\sigma$ -functions were also contracted, the factor being 1.213, while  $2p\pi$  was expanded by a factor 0.932. The ratio between  $2p\sigma:2p\pi$  was 1.3. The present results show much smaller changes from best-atom to best-molecule values. The ratio  $2p\sigma:2p\pi$  was found to be 0.93, i.e. *smaller* than unity for the outermost orbital, the others being unity, cf. Table 1. This result indicates that the higher flexibility of a double-zeta compared to a minimal basis is

sufficient to ensure almost optimal conditions, the optimization of the exponents being rather unimportant in the double-zeta case.

In order to study the level of accuracy obtainable by the optimized (C/9, 5) basis, the results for ethylene were compared with results obtained by both larger and smaller bases. The largest basis used in the present investigation was the best-atom (C/11, 6, 1) basis of van Duynenvelt [4], contracted with the partitionings (6, 2, 1, 1, 1) and (4, 1, 1) to  $\langle C/5, 3, 1 \rangle$ , combined with his (H/5, 1) basis contracted (3, 1, 1) to  $\langle H/3, 1 \rangle$ . This basis is reproduced in Table 1. In view of the experience obtained during optimization of the (C/9, 5) basis no attempt was made to optimize the exponents of this even larger basis. The results emerging from this basis are likely to be very close to the Hartree-Fock limit, cf.  $C_2H_2$  below. Comparison has also been made with previously published results [6], obtained from the (C/7, 3) basis [1], contracted (4, 1, 1, 1), (2, 1) to  $\langle C/4, 2 \rangle$ , combined with Huzinaga's [5] (H/4) basis with a scaling factor 1.25 and contracted (3, 1), to which a  $p$ -function with orbital exponent 0.875 was added to give  $\langle H/2, 1 \rangle$ .

The main part of the calculations were performed at the IBM Research Laboratory, San José, on an IBM 360/91 computer with the program system REFLECT [7]. This program was further developed to include full use of  $C_{2v}$  symmetry both in integral calculation and in SCF symmetry blocking.

### 3. Applications to Ethylene

All the calculations on ethylene were carried out with the following geometry:  $r(CC) = 2.5510$ ,  $r(CH) = 2.0236$  bohrs,  $\angle HCH = 117^\circ$ . The molecule belongs to the point group  $D_{2h}$ . We have chosen a coordinate system with the  $z$ -axis along the CC bond and the  $x$ -axis perpendicular to the molecular plane. With this choice the electronic configuration is  $(1a_g)^2 (1b_{1u})^2 (2a_g)^2 (2b_{1u})^2 (1b_{2u})^2 (3a_g)^2 (1b_{3g})^2 (1b_{3u})^2$ , the last orbital being the  $\pi$ -orbital.

#### 3.1. Ground State Properties: Basis Set Dependency

The different kinds of basis sets employed and the notations used below are collected in Table 2. The basis denoted  $Apd$  is sufficiently accurate to serve as a measure on the goodness of the other basis sets. Unprimed letters  $A, B, C, D$  are used for atom-optimized sets and primed,  $B', D'$ , for ethylene-optimized sets. Lower case letters,  $p, d$ , indicate polarization functions added to the original set.

Calculated values of total energies are listed in Table 3. It is seen that the substantial improvement of 0.07 hartrees in going from the basis  $D$  to  $D'$  is reduced to a minor improvement of 0.01 hartrees of  $B'$  compared to  $B$ . The role of polarization functions however is almost equally important in all three cases  $C, B$  and  $B'$ . Hence these functions do not merely serve as a necessary enlargement of the basis but describe a real polarization though with very low population, ca. 0.01 of an electron at each "polarized" atom. It should also be noted that the basis  $B'pd$  is almost as good as  $Apd$ , cf. also below.

Table 2. Gaussian basis sets

Notation	Optimized for	Size
<i>A</i>	Atoms	(C/11, 6), (H/5)
<i>Ap</i>	Atoms	(C/11, 6), (H/5, 1)
<i>Ad</i>	Atoms	(C/11, 6, 1), (H/5)
<i>Apd</i>	Atoms	(C/11, 6, 1), (H/5, 1)
<i>Bpd</i> , etc.	Atoms	(C/9, 5, 1), (H/4, 1)
<i>B'pd</i> , etc.	Ethylene	(C/9, 5, 1), (H/4, 1)
<i>Cp</i> , etc.	Atoms	(C/7, 3), (H/4, 1)
<i>D</i> , Ref. [3]	Atoms	(C/8, 4), (H/4)
<i>D'</i> , Ref. [3]	Ethylene	(C/8, 4), (H/4)

The orbital energies are very similar for all basis sets. Therefore only the results from *B*, *B'pd* and *Apd* are listed in Table 4. Their connection with ionization energies will be discussed below.

A property which shows a more pronounced sensitivity towards the choice of basis are Mulliken's [8] population numbers. This sensitivity is apparent both in subpopulation and total populations. Some results from the population analysis are displayed in Figs. 1 and 2. Fig. 1 shows gross atomic,  $N(X)$ , net atomic,  $n(X)$ , and overlap populations,  $n(XY)$ , obtained in the various cases.

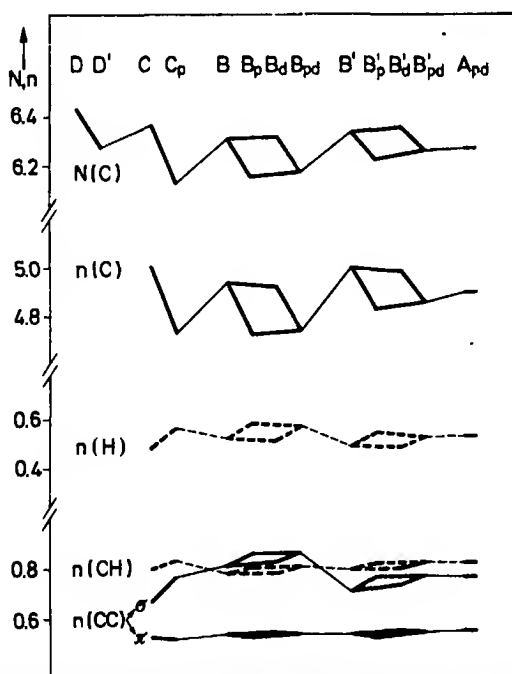


Fig. 1. Total populations of ethylene in various basis sets: gross atomic:  $N(C)$ ; net atomic:  $n(C)$  solid line,  $n(H)$  broken line; overlap  $n(CC)$  solid line,  $n(CH)$  broken line

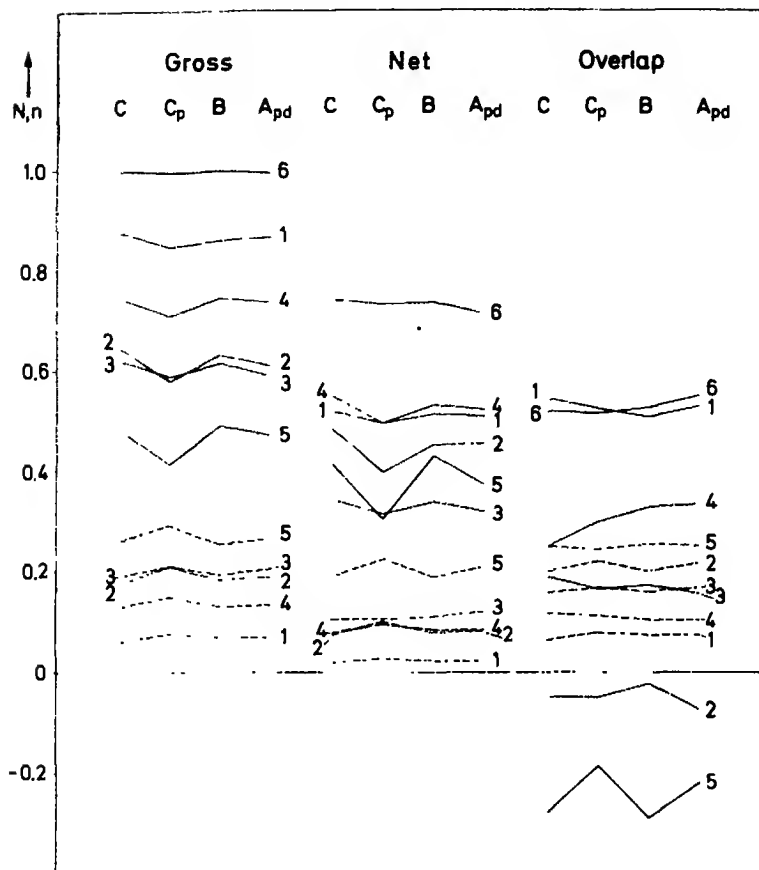


Fig. 2. Orbital populations of ethylene: Solid lines:  $N(C)$ ,  $n(C)$ ,  $n(CC)$ ; broken lines  $N(H)$ ,  $n(H)$ ,  $n(CH)$ . The numbers indicate the various valence orbitals, 6 being the outermost ( $\pi$ ) orbital, cf. Table 4

Comparison of the ethylene-optimized basis  $B'$  to  $B$  without polarization functions only indicates a slightly better agreement with the  $Apd$  basis. The only obvious advantage of  $B'$  over  $B$  is that  $B'$  gives  $n(CH) > n(CC\sigma)$  in accord with the  $Apd$  basis. If  $n(XY)$  can be interpreted as a measure of the strength of the  $XY$  bond, this result is also in agreement with experiments, since energy [9] and force constant [10] of the average  $CH$  bond are slightly larger than those of the  $CC$  single bond.

Polarization functions have similar effects in all the cases,  $C$ ,  $B$  and  $B'$ , cf. Fig. 1. Addition of  $p$  to the hydrogen basis invariably increases not only gross and net atomic populations on hydrogen but also the overlap populations both in  $CH$  and  $CC$  bonds. These  $p$ -functions have a larger effect on all depicted populations than addition of carbon  $d$ -functions. For studies of population-dependent properties an addition of polarization functions only on hydrogen may therefore be of interest. However it must be kept in mind that a formally balanced basis with polarization functions on all atoms, in the present case on



Table 3. Total energy (in hartrees) of ethylene in different basis sets. For notations, see Table 2

Basis	No polariz. function	+ <i>p</i>	+ <i>d</i>	+ <i>p</i> + <i>d</i>
<i>A</i>	—	—	—	—78.0623
<i>B'</i>	—78.0160	—78.0306	—78.0423	—78.0508
<i>B</i>	—78.0060	—78.0287	—78.0327	—78.0471
<i>C</i> [6]	—77.9464	—77.9685	—	—
<i>D'</i> [3]	—77.9083	—	—	—
<i>D</i> [3]	—77.8399	—	—	—

Table 4. Orbital energies (in eV) of ethylene in different basis sets

No.	UAO	<i>D</i> <sub>2h</sub>	Calculation				Exptl. Vertical IP [18]
			— $\epsilon(B)$	— $\epsilon(B'pd)$	— $\epsilon(Apd)$	$\Delta E$ (SCF)	
.	—	1 $a_g$	305.772	305.818	305.625	298.586	
—	—	1 $b_{1u}$	305.730	305.778	305.582	—	
1	2 $s\sigma_g$	2 $a_g$	28.250	27.972	28.066	—	23.7
2	2 $p\sigma_u$	2 $b_{1u}$	21.751	21.698	21.778	20.985	19.2
3	2 $p\pi_y$	1 $b_{2u}$	17.671	17.478	17.586	16.758	15.9
4	3 $d\sigma_g$	3 $a_g$	16.042	15.881	16.011	14.405	14.8
5	3 $d\pi_{yz}$	1 $b_{3g}$	13.925	13.913	14.006	13.313	12.4
5	2 $p\pi_x$	1 $b_{3u}$	10.192	10.054	10.128	8.906	10.5

both hydrogen and carbon, is necessary for a good description of other properties. This is borne out e.g. by Table 3 showing that addition of carbon *d*-functions in all cases means a larger improvement of the total energy than addition of hydrogen *p*-functions. This result is in complete accord with previous studies of the role of polarization functions, see e.g. Millie and Berthier [11].

Figure 2 shows some examples of orbital populations. It is seen that the general picture is the same for all basis sets, but that there are considerable variations concerning the details, much more pronounced than the variations of the orbital energies, cf. Table 4. This underlines the danger of drawing any conclusion from small population differences.

### 3.2. Ionized States of Ethylene

As is well known, orbital energies,  $-\epsilon_i$ , are frequently used as a measure of ionization potentials with reference to Koopmans' theorem. The shortcoming of this measure can be described as rearrangement energy and change in correlation. Although these effects may amount to a few eV's the grouping and the order of the energy levels is in most cases correctly reproduced by the  $\epsilon$ -values. The rearrangement can be accounted for by separate SCF calculations for each ionized level. In the present study such calculations were performed in the *Apd*

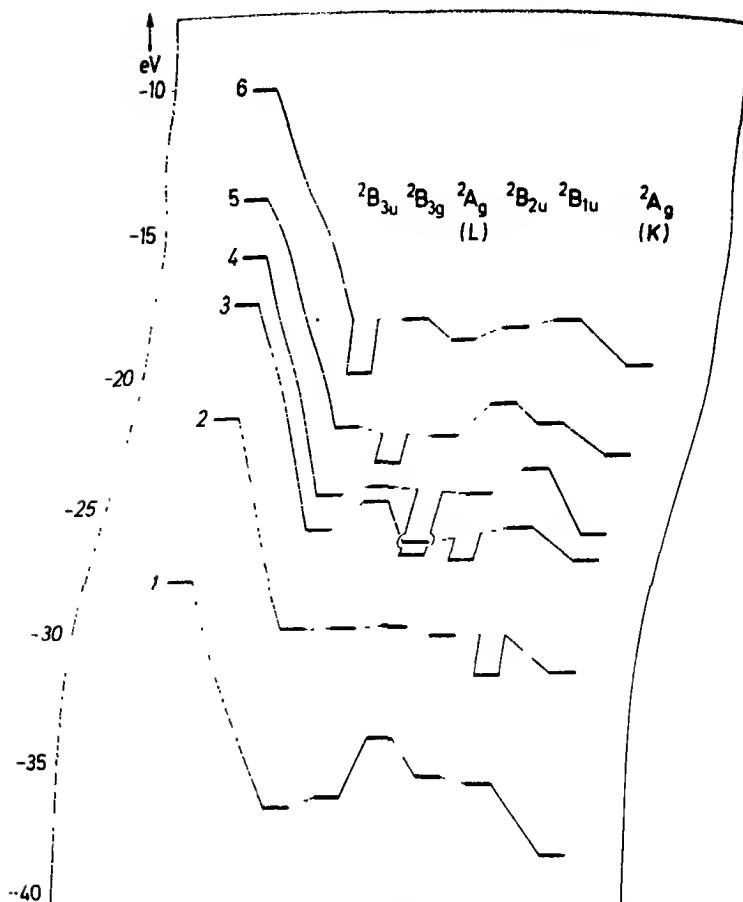


Fig. 3. Values of  $\epsilon$  for ethylene. The numbers refer to the various valence orbitals, cf. Table 4. Values referring to the neutral molecule (to the outermost left) and to various ionized states

basis for the main part of the ionic states that can be described by singly ionized configurations. Ionization potentials,  $\Delta E(\text{SCF})$ , obtained as the difference in total energy between the ion and the neutral molecule are listed in Table 4. In almost all cases the values of  $\Delta E(\text{SCF})$  are closer to the experimental IP values than are the  $-\epsilon_i$ 's.

In spite of the deficiency of the  $\epsilon$ -values it is interesting to see how these quantities will change upon ionization. The  $\epsilon$ -values obtained for the different states are displayed in Fig. 3. It is perspicuous that upon ionization of a valence orbital the  $\epsilon$ -values of all substantially doubly filled orbitals are shifted by almost the same amount, the average shift of a valence orbital being 8.4 eV and of an inner orbital (not displayed in Fig. 3) 9.4 eV. This invariability is particularly interesting, since Koopmans' theorem does not apply to these  $\epsilon$ 's. Using similar assumptions it can be shown that the ionization potential is the sum of the  $\epsilon$ -value and an exchange integral. Experience shows that MO exchange integrals are fairly constant within the valence shell. This indicates the utility of the independent particle model for a first estimate of the energies involved in multiple ionization.

Figure 3 also shows that  $\epsilon$  of the singly occupied orbital is shifted by about 10 eV. Since this  $\epsilon$ -value does submit to Koopmans' theorem, it could be foreseen that at this shift should be different from the other valence orbital shifts. When an inner orbital is ionized (the  $^2A_g$  state furthest to the right in Fig. 3), the average valence orbital shift is 10 eV, whereas the  $1a_g$  orbital itself (not displayed) is shifted by 40 eV. It should be noted that this kind of ionized configuration with retention of the  $g, u$ -restriction for each orbital does not correctly describe the experimental situation of K shell ionization since the physical process leaves a vacancy in a localized orbital, cf. Bagus and Schaefer [12].

Another point to be noted from the values displayed in Fig. 3 is that after ionization the level order remains unchanged except for the  $^2A_g$  state, where the vacancy is in the  $3a_g$  orbital. The coupling between the  $3a_g$  and the  $2a_g$  levels is seen to be exceptionally strong, the pronounced raising of the  $2a_g$  level being counterbalanced by a more than average lowering of the  $3a_g$  level. The result is a reversal of the order between the  $3a_g$  and the adjacent  $1b_{2u}$  level.

#### 4. Applications to Acetylene

A comparison between ethylene and acetylene is clearly of interest for chemical reasons. The comparison is also informative from the technical point of view since McLean and Yoshimine [13] have carried out a very accurate calculation on acetylene, estimated to lie 0.001 hartrees from the Hartree-Fock limit.

All our calculations on acetylene were carried out with the same geometry as adopted by McLean and Yoshimine, i.e.:  $r(\text{CC}) = 2.2810$ ,  $r(\text{CH}) = 2.0020$  bohrs. The electronic configuration of the ground state is  $(1\sigma_g)^2 (1\sigma_u)^2 (2\sigma_g)^2 (2\sigma_u)^2 (3\sigma_g)^2 (\pi_u)^4$ . Eight different basis sets were employed: four of type  $B'$  and four of type  $A$ , cf. Table 2. Some results are listed in Tables 5, 6, 7 and 8.

Total energies are displayed in Table 5. For comparison results from McLean and Yoshimine [13] and Snyder and Basch [14] are also included. Again polarization functions improve the energy substantially (0.03 hartrees) in both basis sets,  $B'$  and  $A$ . It is also seen that the  $Apd$  basis is very close to the Hartree-Fock limit.

Table 6 shows some typical values of orbital energies. The proximity of the  $pd$  basis to the Hartree-Fock limit also in this respect is clear.

Total and orbital populations are listed in Tables 7 and 8. Again the values of these quantities are rather sensitive to the choice of basis. As an example, the

Table 5. Total energy (in hartrees) of acetylene in different basis sets

basis	No polariz. function	+ $p$	+ $d$	+ $p + d$
	-76.8133	-76.8227	-76.8423	-76.8482
	-76.8016	-76.8100	-76.8275	-76.8328
cf. [13]	---	---	---	-76.8540
cf. [14]	-76.7919	---	---	---

Table 6. Orbital energies (in eV) of acetylene in different basis sets

Orbital UAO	Calculation $-\epsilon$					Exptl. Vertical IP [18]
	$B'$	$B'p$	$B'pd$	$Apd$	Ref. [13]	
$1\sigma_g$	306.531	306.569	306.283	305.999	305.958	—
$1\sigma_u$	306.434	306.471	306.186	305.898	305.857	—
$2s\sigma_g$	28.234	28.295	27.909	28.041	28.016	23.4
$2p\sigma_u$	20.777	20.735	20.863	20.949	20.948	18.6
$3d\sigma_g$	18.520	18.473	18.415	18.560	18.579	16.7
$2p\pi_u$	11.178	11.168	11.044	11.148	11.166	11.4

Table 7 Total populations of acetylene in different basis sets. Gross atomic  $N(X)$ , net atomic  $n(X)$  and overlap  $n(XY)$  populations

Population	Ref. [14]	$B'$	$B'p$	$B'pd$	$Apd$
$N(C)$	6.258	6.270	6.201	6.206	6.217
$N(H)$	0.742	0.730	0.799	0.794	0.783
$n(C)$	4.905	4.883	4.794	4.765	4.836
$n(H)$	0.380	0.395	0.439	0.438	0.399
$n(CC)$	1.984	2.110	2.093	2.120	1.992
$n(CH)$	0.753	0.689	0.740	0.741	0.814
$n(\sigma; CC)$	0.860	1.000	0.981	0.953	0.812
$n(\pi, CC)$	1.123	1.116	1.112	1.158	1.180

Table 8. Orbital populations of acetylene

Basis	Population	Orbital			
		$2\sigma_g$	$2\sigma_u$	$3\sigma_g$	$1\pi_u$
$B'$	$N(C)$	0.986	0.657	0.628	1.000
	$N(H)$	0.014	0.343	0.372	—
	$n(C)$	0.626	0.405	0.410	0.721
	$n(H)$	0.010	0.161	0.224	—
	$n(CC)$	0.710	0.140	0.144	0.558
	$n(CH)$	0.016	0.338	0.334	—
$Apd$	$N(C)$	0.955	0.629	0.640	0.996
	$N(H)$	0.045	0.371	0.360	0.004
	$n(C)$	0.565	0.435	0.440	0.698
	$n(H)$	0.014	0.180	0.205	0.000
	$n(CC)$	0.720	-0.002	0.094	0.590
	$n(CH)$	0.050	0.396	0.356	0.006

CC overlap population of the  $2\sigma_u$  orbital is changed from +0.140 in the  $B'$  basis to -0.002 in the  $Apd$  basis. The corresponding values of the  $3\sigma_g$  orbital are +0.144 and +0.094.

### 5. Applications to Benzene

Benzene has frequently been chosen as a test molecule for the evaluation of theoretical methods. Besides, this molecule is of considerable interest by itself. Schulman *et al.* [3] have applied their minimal basis to benzene and also employed an extended basis [15]. An even larger basis has been used by Buenker *et al.* [16] in an *ab initio* study of the benzene spectrum. Very recently, Almlöf *et al.* [17] have performed *ab initio* calculations on azabenzenes, including the unsubstituted parent molecule, employing the above-mentioned (C/7, 3) basis [1].

The present calculations on benzene were carried out with the following geometry:  $r(\text{CC}) = 2.6323$ ,  $r(\text{CH}) = 2.0409$  bohrs. On account of the size of the molecule the choice of basis set had to be much more restrictive than for acetylene and ethylene. Only the bases  $B'$  and  $B'p$  were used. Some results are listed in Tables 9, 10 and 11.

Table 9 presents the total energies obtained. For comparison also some results from previous calculations [3, 15, 16, 17] are included. From Tables 3 and 5 it can be seen that for molecules with two carbon atoms the  $B'$  basis gives rise to total energy values less than 0.06 hartrees, or 0.7% above the Hartree-Fock limit. With the reasonable assumption that the accuracy is comparable in the case of benzene, the total energy of the  $B'p$  basis is likely to be less than 0.15 hartrees or 0.6% above the Hartree-Fock limit, which is estimated to be close to  $-230.82$  hartrees.

Table 9. Total energy (in hartrees) of benzene in different basis sets

Contr. basis	$\{ \langle \text{C}/2, 1 \rangle$ $\langle \text{H}/1 \rangle$	$\langle \text{C}/3, 2 \rangle$ $\langle \text{H}/1 \rangle$	$\langle \text{C}/3, 1 \rangle$ $\langle \text{H}/1 \rangle$	$\langle \text{C}/4, 2 \rangle$ $\langle \text{H}/2, 1 \rangle$	$B'$	$B'p$
Ref.	[3]	[15]	[16]	[17]	Present work	
Energy	-230.318	-230.463	-230.375	-230.476	-230.658	-230.679

Table 10. Orbital energies (in eV) of benzene

Orbital UAO	$m$	$D_{6h}$	Calculation $-\epsilon$		Exptl. Vertical IP [18]
			$B'$	$B'p$	
		$1a_{1g}$	306.11	306.15	—
		$1e_{1u}$	306.10	306.13	—
		$1e_{2g}$	306.06	306.10	—
		$1b_{1u}$	306.05	306.08	—
2s	0	$2a_{1g}$	31.533	31.562	25.8
2p	1	$2e_{1u}$	27.762	27.759	22.7
3d	2	$2e_{2g}$	22.489	22.456	19.2
3s	0	$3a_{1g}$	19.457	19.415	17.0
4f	3	$2b_{1u}$	17.463	17.423	15.5
4f'	3	$1b_{2u}$	16.930	16.904	14.9
3p	1	$3e_{1u}$	16.091	16.079	14.0
2p $\pi$	0	$1a_{2u}$	13.705	13.699	12.5
4d	2	$3e_{2g}$	13.392	13.376	11.8
3d $\pi$	1	$1e_{1g}$	9.199	9.191	9.3

Orbital energies are listed in Table 10. A comparison between calculated orbital energies and vertical ionization potentials has recently been made by Almlöf *et al.* [17]. These authors have suggested an empirical correction of the  $\epsilon$ -values for adjustment to vertical ionization potentials. After addition of the correction terms, they found the following order of the lowest ionization potentials:  $\pi(1e_{1g})$ ,  $\pi(1a_{2u})$ ,  $\sigma(3e_{2g})$ .... Lindholm *et al.* [18], on the other hand, interpret their measurements as supporting the order  $\pi$ ,  $\sigma$ ,  $\pi$ , .... This is also the calculated order both in the present work and in Ref. [17] before correction. The calculated energy difference:  $\epsilon(3e_{2g}) - \epsilon(1a_{2u})$  is found to have almost the same value in the bases  $B'$  (0.31 eV) and  $B'p$  (0.32 eV) as in Ref. [17] (0.36 eV). Therefore the present results cannot contribute any new argument to this controversy.

In the first column of Table 10, the united atom assignments of the various orbitals are listed. The table also includes the  $m$ -values given by Slater [19]. His discussion of the band structure of benzene can now be illustrated both by experimental [18] and calculated values. These values are displayed in Fig. 4. It is seen that the over-all picture is similar to that given by Slater. The virtual orbital energies, included in Fig. 4, are of course an artefact and are strongly basis dependent. It is however interesting that disregarding the highest depicted empty band, both experiments and calculations support Slater's general picture. A closer analysis of the calculated orbitals reveals, however, that the kind of bonding, proposed by Slater for the  $\sigma$ -bands, is not borne out by the calculations. This conclusion is based upon the assumption that overlap populations are of some guidance concerning bonding properties. Although well aware of the weakness of this assumption, we expect it to be valid for the present analysis, *vide infra*. Orbital populations from the  $B'$  basis are collected in Table 11. The total populations from the  $B'p$  basis are also given. The details of the  $B'p$  population are rather similar to those of the  $B'$  basis, the essential difference being a slight increase in some orbital  $n(H)$ -values, particularly in the second band. These values are the source of the increase by 0.062 of the total  $n(H)$  population shown in the

Table 11 (Orbital ( $B'$  basis) and total ( $B'$  and  $B'p$  bases) populations of benzene

Orbital $m$	$D_{oh}$	No.	Populations					
			$n(C)$	$n(H)$	$n(CC)$	$n(C'H)$	$N(C)$	$N(H)$
0	$2a_{1u}$	1	0.162	0.003	0.126	0.011	0.323	0.010
1	$2e_{1u}$	2	0.193	0.007	0.109	0.024	0.313	0.020
2	$2e_{2g}$	3	0.192	0.013	0.083	0.040	0.229	0.034
3	$1h_{2u}$	6	0.225	0.000	0.129	0.000	0.333	0.000
0	$3a_{1g}$	4	0.126	0.049	0.043	0.091	0.233	0.100
1	$3e_{1u}$	7	0.163	0.069	0.065	0.076	0.226	0.108
2	$3e_{2g}$	9	0.221	0.077	0.003	0.093	0.235	0.099
3	$2h_{1u}$	5	0.174	0.073	-0.035	0.155	0.199	0.135
0	$1a_{2u}$	8	0.191	—	0.112	—	0.333	—
1	$1e_{1g}$	10	0.275	—	0.095	—	0.333	—
Total $B'$			4.963	0.458	1.088	0.722	6.234	0.766
Total $B'p$			4.850	0.520	1.115	0.751	6.163	0.837

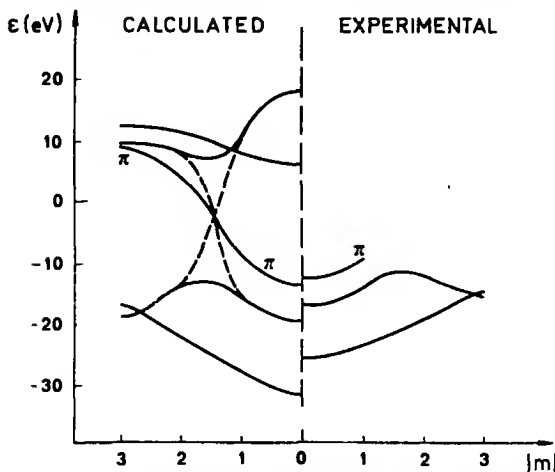


Fig. 4. Energy "bands" of benzene obtained by plotting  $\epsilon$  against (cylindrical)  $|m|$ -values. All bands but one are of  $\sigma$ -type

table. In particular, the orbital  $n(\text{CC})$  and  $n(\text{CH})$  values were almost identical in the two bases  $B'$  and  $B'p$ . Table 11 shows that  $n(\text{CC})$  is considerable for the first valence band whereas  $n(\text{CH})$  is very small. This is valid for the whole band if we assign the  $1b_{2u}$  orbital to the first band. However according to Table 10 we have found  $\epsilon(1b_{2u}) > \epsilon(2b_{1u})$ . This fact seems to invalidate our assignment of  $1b_{2u}$  to the lowest valence band. On the other hand, experiments [18] show that these two orbitals are almost (accidentally) degenerate. Moreover, calculations on the azabenzenes, see e.g. Ref. [17], on toluene, fluorobenzene, etc., in all cases to our knowledge have given  $\epsilon(1b_{2u}) < \epsilon(2b_{1u})$  in support of our assignment. We therefore conclude that the lowest valence band is CC-bonding. Table 11 also shows that the next valence band is mainly CH-bonding, although the first two orbitals are somewhat CC-bonding as well. The third ( $\pi$ ) band is obviously CC-bonding. The advantage of grouping the orbitals into bands is indicated by these population numbers. An arrangement of the orbitals according to increasing  $\epsilon$ -values (sequential numbers in the third column of Table 11) would have obscured the regularity of the overlap populations. A similar regularity is brought out by the orbital contour maps, cf. Ref. [17]. It is found that all orbitals belonging to the first band have extreme values in CC-regions (atomic and/or bond regions) whereas those of the second band have nodal surfaces close to the carbon nuclei and extreme values in the CH bond directions. Hence these contour maps strongly support our conclusions about bonding properties. In fact, if evidence from contour maps were in contradiction to population numbers we expect contour maps to be the better representatives of the physical situation.

It should be pointed out that the present conclusions concerning bonding and antibonding properties are only in partial agreement with recent publications, e.g. Jonsson and Lindholm [20]. As an example, these authors classify the  $2e_{2g}$  orbital as weakly CC-antibonding. The reason seems to be that one orbital in this pair of degenerate orbitals can be represented by an orbital with a nodal

surface cutting two CC-bonds. However the population numbers in Table 11, +0.083 for  $2e_{2g}$ , are the *average* populations in an orbital from the degenerate pair, one contribution being positive and the other negative. Population numbers as well as contour maps show that the positive contribution will dominate for  $2e_{2g}$  as well as for e.g.  $2e_{1u}$ . Hence both should be classified as CC-bonding. Similarly, the pairs  $3e_{1u}$  and  $3e_{2g}$  are about equally strongly CH-bonding.

## 6. Corresponding Orbitals

Since results for the various molecules have been obtained with the same basis set *B* it is interesting to compare the results. For this purpose we shall give some comments on corresponding orbitals of acetylene and ethylene. As a third member of the series we have chosen  $N_2$ , for which results from a (N/9, 5) basis are available [21].

Corresponding molecular orbitals within a series with the same number of heavy atoms are those with the same UAO (United Atom Orbital) description. Extension to cases with different numbers of heavy atoms can be made. Since there is no unique way of doing this we shall not include benzene in our discussion, although an interesting parallelism between orbitals of  $C_2H_4$  and  $C_6H_6$  can be found. Correlation diagrams between corresponding orbital energies in the series  $C_2$  to  $C_2H_6$  have been published by Buenker *et al.* [22]. We shall therefore limit ourselves to a discussion of bonding properties with reference to overlap populations and contour maps. Overlap populations are collected in Table 12. The  $2s$  and  $2p\pi'$  orbitals are clearly XX bonding within the series. The  $2p\pi$  orbital (in the  $C_2H_4$  plane) is also bonding although in  $C_2H_4$  the bonding includes both CC and CH bonds. These conclusions are nicely borne out by the contour maps of Fig. 5. From the map of the  $2p\pi$  orbital in  $C_2H_4$  it is easy to predict that the  $2p\pi(1b_{2u})$  orbital of  $C_2H_4$  must be not only CC- but also CH-bonding, since this bond is lying in a region which is strongly bonding even before distortion by the addition of a new proton. The character of the  $2p\sigma$  and  $3p\sigma$  orbitals are not so easily interpreted from the population numbers. These numbers indicate that  $2p\sigma$  should be strongly antibonding in  $N_2$ , weakly

Table 12. Overlap populations of corresponding valence orbitals

UAO	$n(XY)$	Molecule		
		$N_2$ [21]	$C_2H_2$	$C_2H_4$
$2s$	$n(XX)$	0.77	0.71	0.51
	$n(XH)$	—	0.01	0.08
$2p\sigma$	$n(XX)$	-0.58	0.14	0.02
	$n(XH)$	—	0.34	0.20
$2p\pi$	$n(XX)$	0.49	0.56	0.17
	$n(XH)$	—	—	0.16
$2p\pi'$	$n(XX)$	0.49	0.56	0.53
	$n(XH)$	—	—	—
$3d\sigma$	$n(XX)$	-0.01	0.14	0.33
	$n(XH)$	—	0.34	0.11



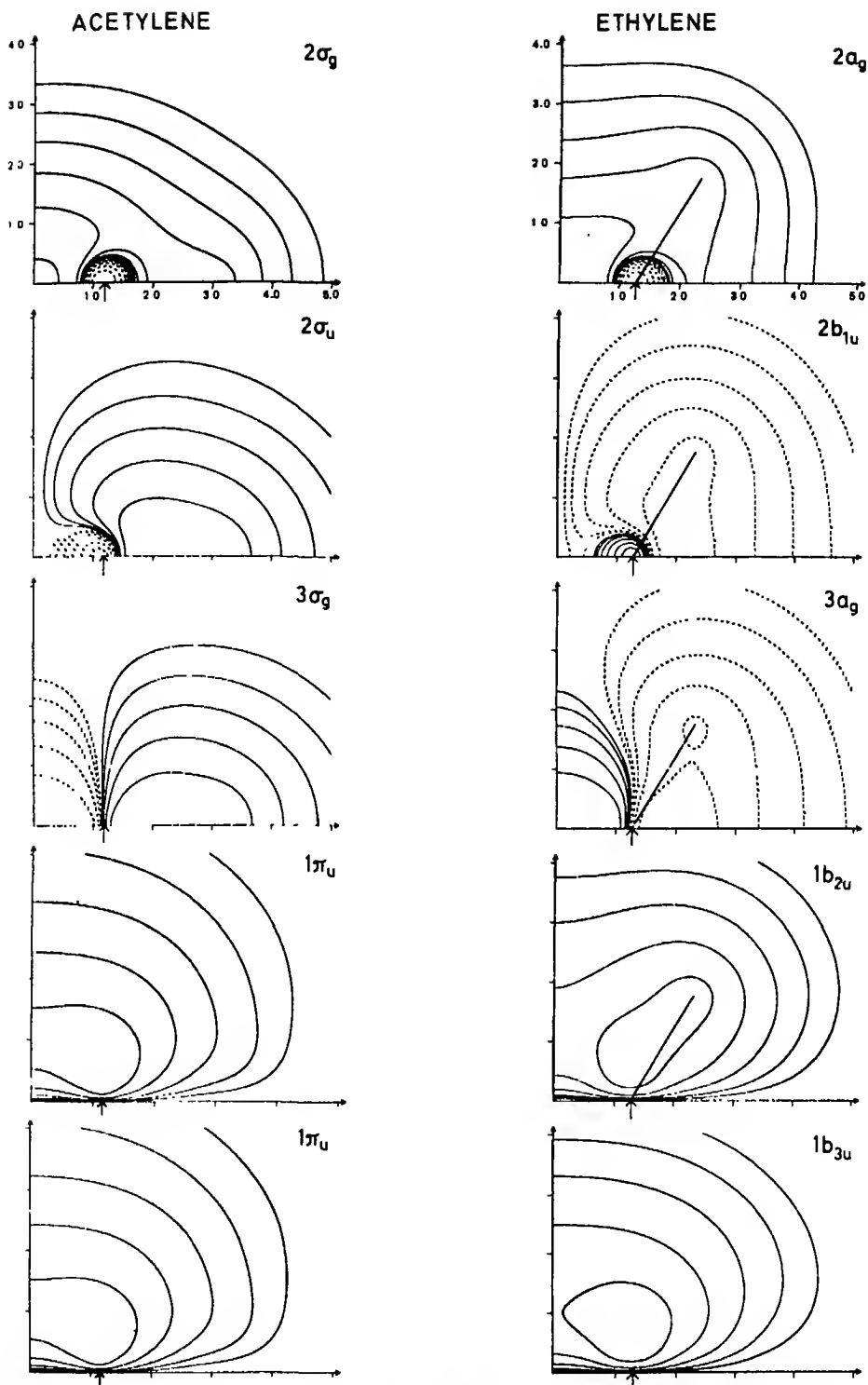


Fig. 5. Contour maps in the molecular plane of pairs of corresponding orbitals. Regions with different signs are indicated by solid and broken lines respectively. The arrows point towards the position of the carbon nucleus on the z-axis. The solid line in ethylene maps is drawn along the CH bond

CC-bonding in  $C_2H_2$  and CC-nonbonding in  $C_2H_4$ . Similarly the population numbers of  $3d\sigma$  indicate this orbital to be non-bonding in  $N_2$ , slightly CC-bonding in  $C_2H_2$  and more CC-bonding in  $C_2H_4$ . Moreover, according to these numbers the  $2p\sigma(2\sigma_u)$  and  $3d\sigma(3\sigma_g)$  orbitals of acetylene should be almost identical. However the contour maps of Fig. 5 strongly indicate that  $2p\sigma$  is XX-anti-bonding and  $3d\sigma$  XX-bonding throughout the series. In particular, the difference between  $2\sigma_u$  and  $3\sigma_g$  of  $C_2H_2$  is clearly demonstrated. These results underline the precautions that must be taken at the interpretation of population numbers. In this context it might be mentioned that we have tried to find a correlation between trends in force constants given by trends in vibration frequencies and population numbers. No simple picture emerged from these studies. Our experience points rather towards contour maps as a cheap and quite useful tool for the discussion of bonding properties. We are well aware that these maps cannot give definite answers in all cases and that more elaborate analysis as given e.g. by Buenker *et al.* [22] and by Clementi and Popkie [23] may be useful and sometimes necessary. It is however interesting that our conclusions from contour maps concerning  $C_2H_2$  and  $C_2H_4$  are in full accord with the extensive calculations by Clementi and Popkie [23].

## 7. Conclusions

The results presented in the preceding Sections show that the "double zeta" basis  $B'$  is sufficiently large to give adequate results for several properties of physical interest. However it is also clear that the non-optimized basis  $B$  could have been used with practically the same results. Even the smaller  $(C/7, 3)$  basis set is seen to be in accord with the  $B'$  basis when double-zeta contraction is employed.

As pointed out many times before we have found that population analysis is so strongly basis set dependent that erroneous conclusions might emerge from an improvident application. On the other hand, contour maps seem to be a very useful instrument, both less sensitive to the choice of basis set and more informative about details of the electronic density distribution. The main drawback at present may be the difficulty to publish this information in an inexpensive way.

The usefulness of the corresponding orbital concept has been demonstrated, and might have been pursued further. It seems however to be most efficient for series of about equally sized molecules.

In the case of benzene it has been shown that Slater's energy-band plot is rather useful for classifying bonding properties.

*Acknowledgments.* We are greatly indebted to IBM World Trade Corporation and to the Swedish IBM Corporation for granting a fellowship to one of us (P.S.) that made it possible to perform the larger part of the calculations at IBM Research Laboratory in San José. Part of this work has been supported by a grant from the Swedish Board for Technical Development and the Swedish Natural Science Research Council. We are grateful to the members of the Stockholm Quantum Chemistry group for unpublished details of their calculations. Thanks are due to Professor E. Lindholm and coworkers for preprints prior to publication.

## References

1. Roos, B., Siegbahn, P.: Theoret. Chim. Acta (Berl.) 17, 209 (1970)
2. Roos, B., Siegbahn, P.: Theoret. Chim. Acta (Berl.) 17, 199 (1970)
3. Schulman, J. M., Hornback, C. J., Moskowitz, J. W.: Chem. Physics Letters 8, 361 (1971)
4. van Duynenvelt, F.: Private communication
5. Huzinaga, S.: J. Chem. Phys. 42, 1293 (1965)
6. Meza, S., Wahlgren, U.: Theoret. Chim. Acta (Berl.) 21, 323 (1971)
7. Siegbahn, P.: Chem. Physics Letters 8, 245 (1971)
8. Mulliken, R. S.: J. Chem. Phys. 23, 1833 (1955)
9. See e.g. Cottrell, T. L.: The strengths of chemical bonds. London: Butterworths 1954
10. See e.g. Herzberg, G.: Molecular spectra and molecular structure Vol. II. Infrared and raman spectra of polyatomic molecules. New York: Van Nostrand 1951
11. Millie, P., Berthier, G.: Intern. J. quant. Chem. 2 S, 67 (1968)
12. Bagus, P. S., Schaefer III, H. F.: J. Chem. Phys. 56, 224 (1972)
13. McLean, A. D., Yoshimine, M.: Tables of linear molecular wave functions, Supplement to IBM J. Res. Develop. 12, 206 (1968)
14. Snyder, L. C., Basch, H.: Molecular wave functions and properties. Wiley-Interscience 1972
15. Schulman, J. M., Moskowitz, J. W.: J. Chem. Phys. 47, 3491 (1967)
16. Buenker, R. J., Whitten, J. L., Petke, J. D.: J. Chem. Phys. 49, 2261 (1968)
17. Almlöf, J., Johansen, H., Roos, B., Wahlgren, U.: J. Electron Spectr. In press
18. Fridh, C., Åsbrink, L., Lindholm, E.: Chem. Phys. Letters 15, 282 (1972)  
Lindholm, E., Fridh, C., Åsbrink, L.: Discussions Faraday Soc. 54, 127 (1972)  
Lindholm, E.: Discussions Faraday Soc. 54 (1972)  
Price, W. C.: Discussions Faraday Soc. 54 (1972)
19. Slater, J. C.: Quantum theory of molecules and solids I. New York: McGraw-Hill 1963
20. Jonsson, B. Ö., Lindholm, E.: Arkiv Fysik 39, 65 (1969)
21. Roos, B.: Private communication
22. Buenker, R. J., Peyerimhoff, S. D., Whitten, J. L.: J. Chem. Phys. 46, 2029 (1967)
23. Clementi, E., Popkie, H.: J. Chem. Phys. 57, 4870 (1972)

Prof. Inga Fischer-Hjalmars  
Institute of Theoretical Physics  
University of Stockholm  
Vanadisvägen 9  
S-11346 Stockholm, Sweden



## Complementary Molecular Orbital Investigations on the Conformation of Choline Derivatives\*

Bernard Pullman and Philippe Courrière

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique associé au C.N.R.S.  
13, rue P. et M. Curie, Paris 5è

Received March 12, 1973

Quantum-mechanical computations by the PCILO method, applied previously to the study of the conformational properties of acetylcholine and its derivatives modified in the central part of this molecule, are extended to modifications involving its cationic head and its ester terminal. The replacement of the methyl groups of the cationic head by hydrogens or ethyl groups leads to a steep decline in parasympathomimetic activity. It is shown that the triethyl derivative conserves the *gauche* form as the most stable one. The redistribution of the electronic charges at the onium group implies, however, a transition from an ionic to a hydrophobic binding. The replacement of the methyls by two or three hydrogens leads to a different preferred *gauche-gauche* conformation. The replacement of the methyl group at the ester terminal by a phenyl ring enables a comparison with the conformational properties of local anesthetics. The study brings about evidence, substantiated by NMR spectroscopy, that acetylcholine analogs and protonated local anesthetics are conformationally similar. Choline ethers also show a general preference for a *gauche* conformation. Nevertheless, biological studies do not indicate a constant correlation between conformation and biological potency. Conformational analogies or discrepancies alone cannot thus account for the fine details of the biological activity which must depend also on the electronic structure.

**Key words:** PCILO method · Acetylcholine, conformation of ~ Acetylcholine, structure-activity relation of ~ Local anesthetics.

### Introduction

The conformational properties of choline derivatives, in particular those of acetylcholine are among the most extensively investigated problems in molecular pharmacology, both theoretically and experimentally (for general reviews see e.g. [1–4]). In our previous theoretical work in this field we have centered our attention essentially on acetylcholine itself and on the derivatives obtained from it by a modification of the central part of this molecule: substitution of methyl groups in place of the hydrogens at carbons  $\alpha$  and  $\beta$ , replacement of the ester oxygen by sulfur, etc. This paper presents an extension of this study to the effect of modifications involving the two ends of this fundamental compound: the cationic head and the ester group, considered both generally as the principal sites of interaction of the drug with the cellular receptor (see e.g. [5, 6]); with the onium group essential for the intrinsic activity and the ester group contributing substantially to the affinity [7].

\* This work was supported by the A.T.P. N° A 655-2303 of the C.N.R.S.

### The Method

The method used here is the PCILO (Perturbative Configuration Interaction using Localized Orbitals) procedure utilized also in the preceding papers of this series [1, 2]. The geometrical input data correspond, unless otherwise stated, to the crystal structure of acetylcholine found in its chloride [8] with standard variations introduced for the modified portion of the analogs studied. We remind that the fundamental torsion angles investigated are  $\tau_1$  and  $\tau_2$  (as indicated in I) defined as  $\tau_1 = \text{C}_6-\text{O}_1-\text{C}_5-\text{C}_4$  and  $\tau_2 = \text{O}_1-\text{C}_5-\text{C}_4-\text{N}^+$ .  $\tau_0 = \text{R}_4-\text{C}_6-\text{O}_1-\text{C}_5$  and  $\tau_3 = \text{C}_5-\text{C}_4-\text{N}^+-\text{R}_1$  are fixed following standard stereochemical considerations and numerous experimental evidence at  $180^\circ$ . Following the usual convention, a torsion angle  $\tau$  of the bonded atoms  $\text{A}-\text{B}-\text{C}-\text{D}$  is the angle between the planes ABC and BCD. Viewed from the direction of A,  $\tau$  is positive for clockwise and negative for counterclockwise rotations. The value  $\tau = 0^\circ$  corresponds to the planar-*cis* arrangement of the bonds AB and CD.

### Results and Discussion

#### A. Modification of the Cationic Head

The most active compounds in the parasympathomimetic series contain the  $-\text{N}^+(\text{C}_2\text{H}_5)_3$  group as cationic head. Successive replacement of the methyl groups by either hydrogen or ethyl leads to a steep decline in parasympathomimetic activity [7, 9, 10]. Because of the tendency of many authors to associate this activity, at least in part, with the conformational properties of molecules of this series, it seems obviously useful to investigate the influence of these structural modifications upon the conformation of acetylcholine.

Figure 1 represents the results of computations for the triethyl analog of acetylcholine (I,  $\text{R}_1 = \text{R}_2 = \text{R}_3 = \text{C}_2\text{H}_5$ ,  $\text{R}_4 = \text{CH}_3$ ). The conformational energy map appears very similar to that of acetylcholine [1]. In particular the doubly-degenerated global energy minimum corresponds to a *gauche* arrangement of the  $\text{N}^+$  and ester O atoms ( $\tau_2 = \pm 60^\circ$ ,  $\tau_1 = 180^\circ$ ). The extended form (at  $\tau_1 = \tau_2 = 180^\circ$ ) represents a local energy minimum 3 kcal/mole above the global one.

This result must be considered as particularly significant, in particular because it was presumed by some authors [4] that the presence of larger alkyl groups on  $\text{N}^+$  would diminish the tendency of the molecule to adopt a *gauche* conformation. Crystallographic evidence, available essentially in the series of carbamoylcholines, confirms that it need not be so: thus e.g. 2-N, N-diethyl-N-benzylammoniummethylcarbamate bromide II exists in the crystal in the *gauche* conformation, as does also 2-N, N-dimethyl-N-ethyl-ammoniummethylcarbamate III [11, 12].

It is therefore obvious that the decrease of parasympathomimetic activity in the  $\text{N}^+$ -triethyl derivative cannot be ascribed to a changement in molecular conformation. On the other hand it may reasonably be attributed, at least in part, to: 1) the modification of the dimensions of the cationic head and 2) the modification in the electronic properties of the cationic head; both may profoundly

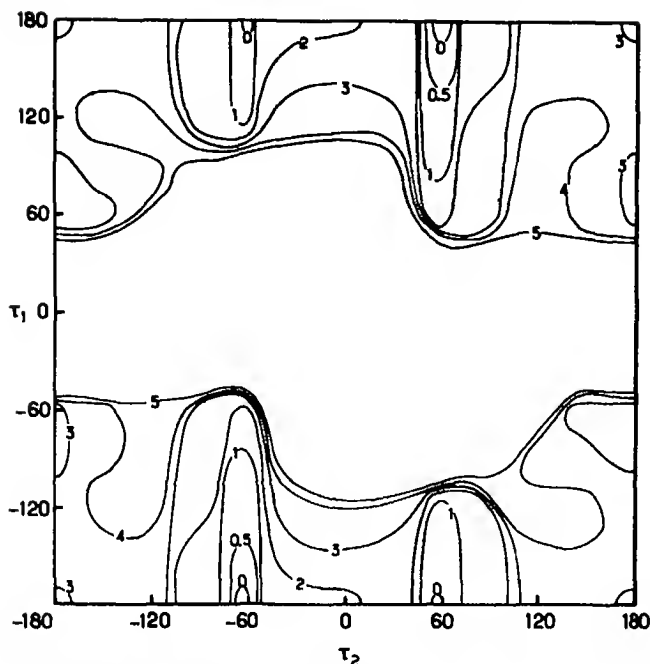


Fig. 1. Conformational energy map for compound 1,  $R_1=R_2=R_3=C_2H_5$ ,  $R_4=CH_3$ . Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

perturb or even preclude the interaction of this head with the anionic receptor site. The role of the possible modification of the dimensions being straightforward we shall center our attention on the nature of the possible modifications in the electronic properties.

As indicated already in Ref. [1], one of the most striking theoretical results on the electronic structure of the cationic head of acetylcholine was to point out that the net positive charge, which in the usual chemical representation<sup>1</sup> is localized on the quaternary N atom, is in fact distributed among the adjacent methyl and methylene groups, leaving the "N<sup>+</sup>" atom nearly neutral (Fig. 2a). We are considering here, of course, the *total* net electronic charges, i.e. a summation of the net  $\sigma$  and  $\pi$ -charges, where "net" charges denote an excess or deficit at each atom in relation to the number of electrons the atom would possess in an isolated state. Thus in acetylcholine 70% of the net positive charge is distributed among the three attached methyl groups, essentially among their hydrogens which carry thus each about 0.07 positive electronic charge as opposed to the usual positive charge of 0.04 e found generally on H atoms linked to saturated carbons. These three methyl groups thus form a large ball of spread-out positive electricity to which the designation of a hydrophylic cationic center is appropriate. The remaining fraction of the positive charge seems to be concentrated on the two

<sup>1</sup> And in some simple quantum-mechanical computations: see e.g. Zull, J.E., Hopfinger, A.J.: *Science* **165**, 512 (1969).

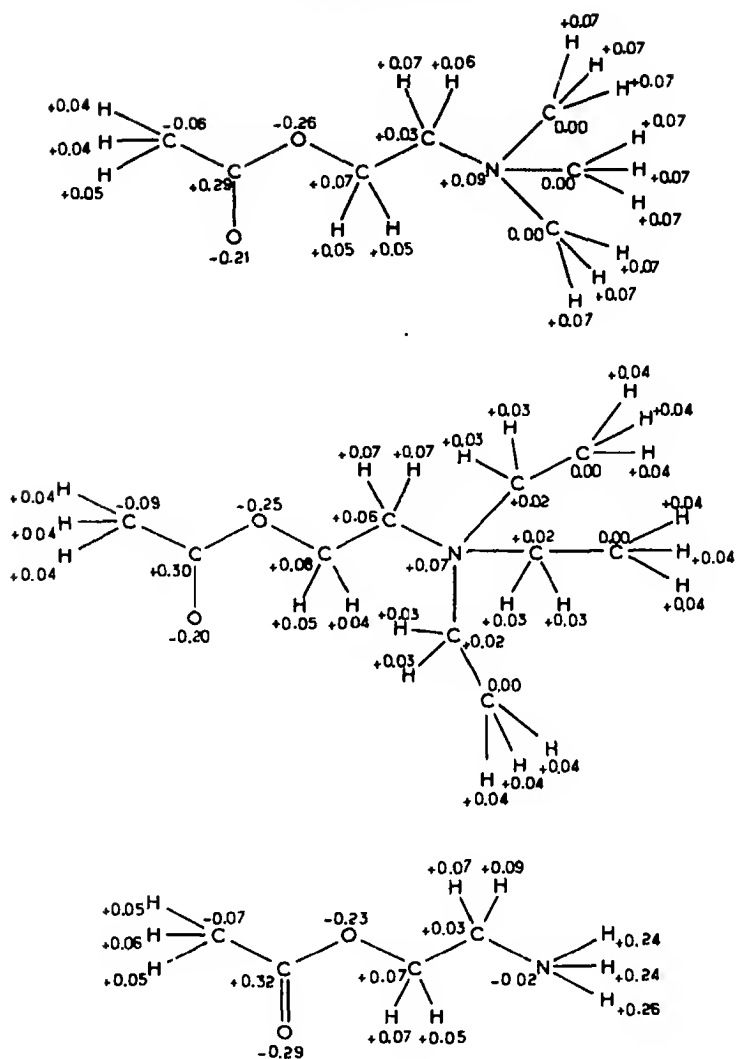


Fig. 2. Distribution of net electronic charges in: a) acetylcholine, b) *I*,  $R_1 = R_2 = R_3 = C_2H_4$ ,  $R_4 = CH_3$  and c) *I*,  $R_1 = R_2 = R_3 = H$ ,  $R_4 = CH_3$

$CH_2$  groups of the backbone of acetylcholine, further enlarging the dimension of the cationic moiety of this molecule.

We may now consider the electronic state of the cationic head in the  $-N^+(C_2H_5)_3$  analog of acetylcholine. This is shown in Fig. 2b. A drastic change is observed with respect to the situation in acetylcholine. The excess positive charge of the cationic head is now shared by a substantially increased number of hydrogen atoms with the result that the net positive charge carried by each of them (0.03 e on the hydrogens of the  $CH_2$  groups and 0.04 e on the hydrogens of the  $CH_3$  groups of the ethyl substituents) is now of the order of magnitude of the



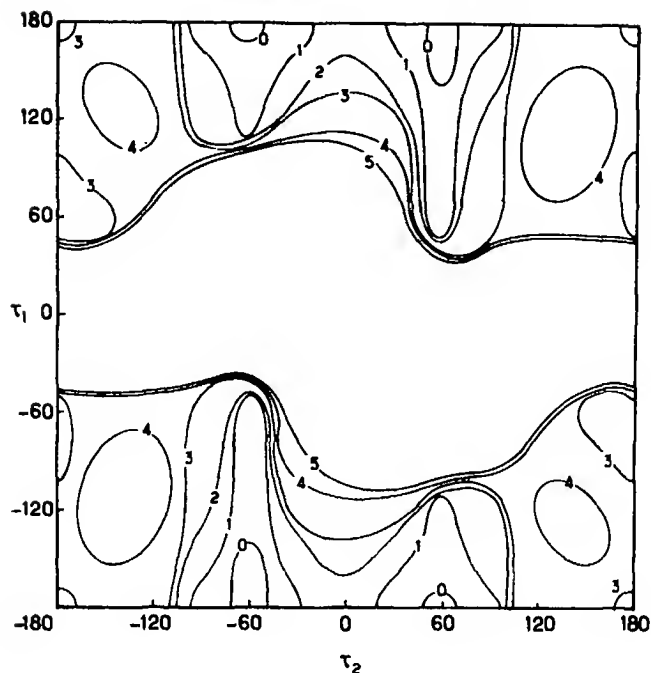


Fig. 3. Conformational energy map for compound  $1$ ,  $R_1 = R_3 = \text{CH}_3$ ,  $R_2 = \text{H}$ ,  $R_4 = \text{CH}_3$ . Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

charge carried usually by hydrogens attached to saturated carbons (it can be seen that the charge of the terminal H atoms of the onium head is identical to the charge carried by the hydrogens of the methyl group of the ester terminal). These hydrogens lose thus their specific character, a transformation which implies a profound modification of the nature of the interaction of the onium group with its surroundings and with a potential receptor. It is to this situation that may probably be ascribed the transition from ionic to hydrophobic binding characteristic of the behaviour of tetraalkylammonium ions upon the increase in size of the N-alkyl substituents and particularly visible upon the replacement of methyl substituents by ethyl ones [10].

We now turn over to the study of the effect of decreasing the size of the cationic head, by replacing its methyl groups by hydrogen atoms, upon the conformational characteristics of the acetylcholine skeleton. Figures 3, 4 and 5 present the conformational energy maps of acetylcholine derivatives in which respectively one, two and three methyl groups of the cationic head have been replaced by hydrogen atoms. In Figs. 3 and 4, the results refer to the "symmetrical" structure of the onium group, i.e., one in which its single hydrogen or methyl carbon is considered to be in the plane of the  $\text{C}_5\text{—C}_4\text{—N}^+$  atoms, this arrangement being found to be more stable than the "non symmetrical" one.

The results indicate that while the replacement of one methyl group of the cationic head by a hydrogen atom does not bring about any major modification

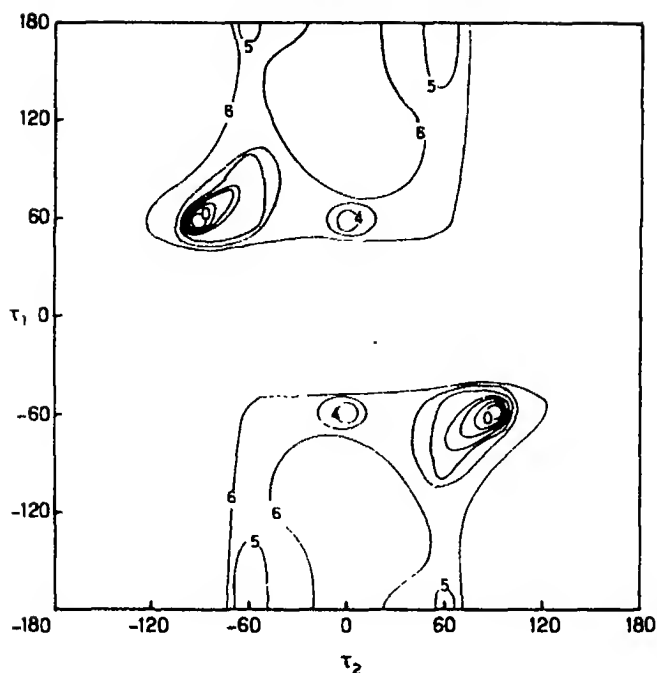


Fig. 4 Conformational energy map for compound I.  $R_1 = R_3 = \text{H}$ ,  $R_2 = \text{CH}_3$ ,  $R_4 = \text{CH}_3$ . Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

of the conformational energy map in comparison to that of acetylcholine and leaves the *gauche* form at  $\tau_1 = 180^\circ$ ,  $\tau_2 = 60^\circ$  as the most stable one, the replacement of two such methyl groups by hydrogens produces a drastic change of the map. The global minimum is transferred to the values of  $\tau_1 = 60^\circ$ ,  $\tau_2 = -90^\circ$  (and the symmetrical values  $\tau_1 = -60^\circ$ ,  $\tau_2 = 90^\circ$ ) which represent a *gauche-gauche* structure with respect to the torsions around both the  $\text{C}_5\text{--C}_4$  and  $\text{O}_1\text{--C}_5$  bonds. The *gauche* conformation characteristic of acetylcholine is now a local minimum, 5 kcal/mole above the global one. The fully extended form, which is at 3 kcal/mole above the global minimum in Fig. 3, is at 7 kcal/mole above the global minimum in Fig. 4. Moreover, the large zone between the *gauche* and the *trans* conformations ( $\tau_1 \approx 180^\circ$ ,  $\tau_2 = 90^\circ - 180^\circ$ ), in which a number of investigators place the conformations related to the muscarinic and nicotinic activity (see e.g. Ref. [3]) is in Fig. 4 at a relatively high energy level, 7–8 kcal/mole above the minimum.

The same general situation is observed in Fig. 5 representing the effect of the complete replacement of the methyl groups of the cationic head of acetylcholine by hydrogens, leading to acetyethanolamine.

In this case thus we observe, at least for the compounds represented by Figs. 4 and 5, the triple effect of modifying simultaneously the dimension of the cationic head, the preferred conformation of the whole molecule and the general morphology of important parts of the conformational energy map. It may also be expected that the direct presence of the hydrogen atoms on the quaternary

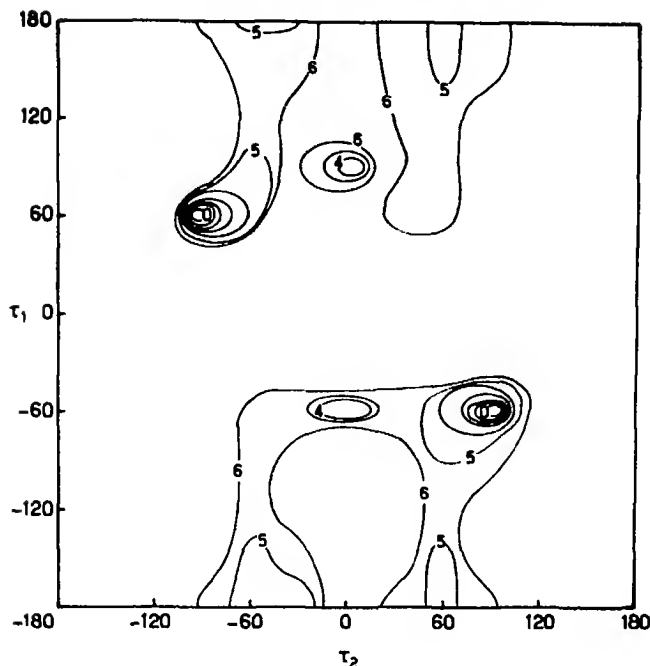


Fig. 5. Conformational energy map for compound I,  $R_1 = R_2 = R_3 = H$ ,  $R_4 = CH_3$ . Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

nitrogen will make these compounds more susceptible to environmental effects in the crystal and in solution. The more so as these hydrogens carry now (as can be seen from Fig. 2c) a very high net positive charge, corresponding, in fact, for each of them to the total net positive charge of the methyl groups that they replace (0.24 e per H atom), the quaternary nitrogen itself becoming even slightly negativ. This complex situation makes it difficult to ascertain the reasons responsible for the decrease of the parasymphomimetic activity in these derivatives. The more precise knowledge of the nature of the different effects linked to the replacement of the  $CH_3$  groups by H atoms should, however, help to determine the biological significance of each of them.

### B. Modification at the Ester Terminal

It seems *a priori* probable that the modification of the molecular structure of acetylcholine at its ester terminal will have a more moderate influence on the conformation of this molecule than the modification of the cationic head. Even moderate modifications of the conformational energy map may, however, be of non negligible significance. We have already discussed [2, 13] the case of carbamoylcholine, a derivative of acetylcholine in which the methyl group of the acetyl fragment is replaced by an amino group ( $I$ ,  $R_1 = R_2 = R_3 = CH_3$ ,  $R_4 = NH_2$ ). This compound exists in the *trans* conformation in the crystalline state

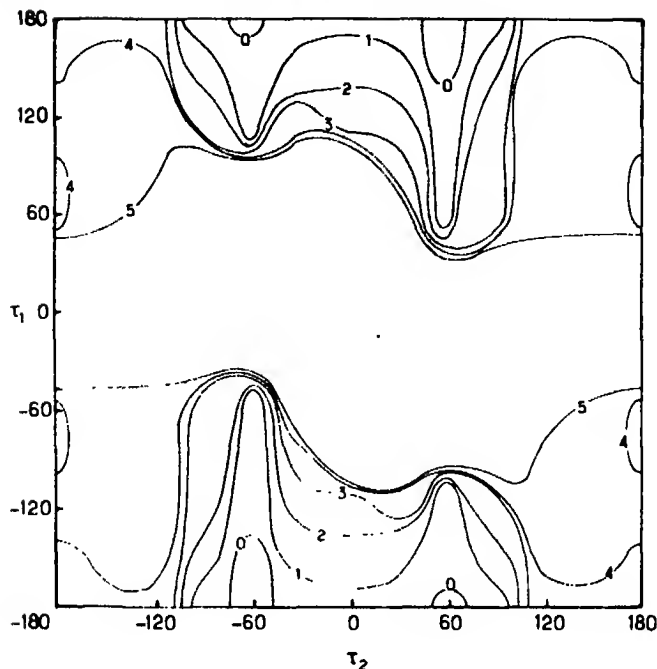


Fig. 6 Conformational energy map for compound IV. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

( $\tau_1 = \tau_2 = 180^\circ$ ) [14], although some of its derivatives substituted by more complex groups at the quaternary nitrogen exist as we have seen in *gauche* forms in their crystals. What is also particularly interesting in connection with carbamoylcholine is that it exists in the *gauche* conformation, similar to that of acetylcholine, in solution [15]. The PCILO conformational energy map of carbamoylcholine [2] shows results which are manifestly relevant to the above enumerated experimental findings. Thus the map, while similar in its overall aspect to that of acetylcholine, contains three practically equivalent global energy minima of which two represent *trans* forms (at  $\tau_1 = 180^\circ$ ,  $\tau_2 = 180^\circ$  and at  $\tau_1 = 80^\circ$ ,  $\tau_2 = -140^\circ - 180^\circ$ ) and one the *gauche* form ( $\tau_1 = 180^\circ$ ,  $\tau_2 = 80^\circ$ ). Carbamoylcholine itself occupies one of the *trans* minima, associated with the most extended form. Other carbamoylcholine derivatives are in conformations which correspond to the two remaining energy minima. Moreover, while the barrier between the *gauche* and *trans* forms is of 4 kcal/mole in acetylcholine it is only of 1 kcal/mole in carbamoylcholine, a situation which is in obvious relation to its relatively easy transition to the *gauche* form observed in solution.

Here we wish to study, in the first place, the simplest derivatives of acetylcholine modified at the ester terminal, namely its nearest homologues, the formic and propionic esters, IV and V. It is known that the parasympathomimetic activity of both these neighbours is appreciably reduced with respect to that of acetylcholine [7, 9].

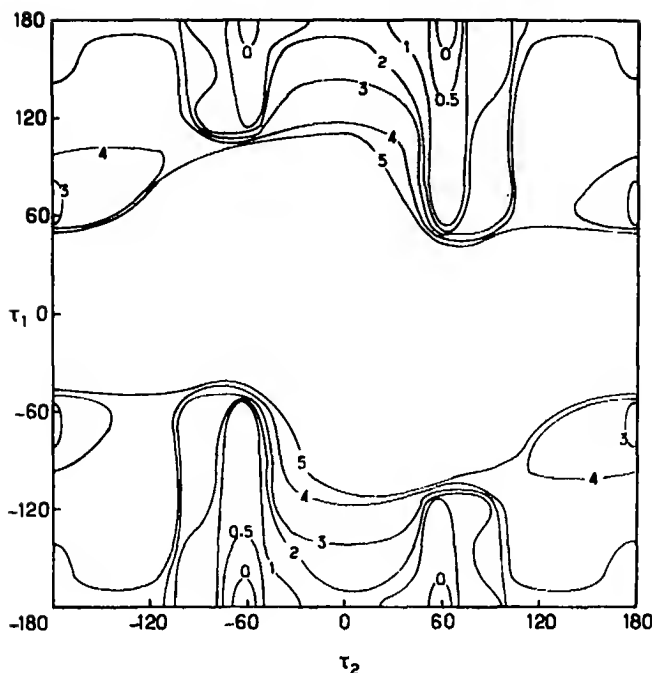


Fig. 7. Conformational energy map for compound V. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

The conformational energy maps of these two derivatives are presented in Figs. 6 and 7. They are very similar and present both a global energy minimum corresponding to the fundamental *gauche* ( $\tau_1 = 180^\circ$ ,  $\tau_2 = 60^\circ$ ) conformation of acetylcholine. They differ, however, both from acetylcholine by being devoid of a local energy minimum for the *trans* conformation ( $\tau_1 = \tau_2 = 180^\circ$ ).

Secondly, we wish to study more particularly the group of 2-dialkylaminoethylbenzoates, compounds which derive from acetylcholine by the replacement of the methyl group of the acetyl fragment by a phenyl ring (and a simultaneous suppression of one of the  $\text{CH}_3$  groups of the cationic head; these molecules may naturally, also be considered in the quaternary state ( $\text{R}_1 = \text{R}_3 = \text{CH}_3$ ,  $\text{R}_2 = \text{H}$ ,  $\text{R}_4 = \text{C}_6\text{H}_5$ ). The introduction of the phenyl group at the esteric end of acetylcholine results in the decrease of the intrinsic parasympathomimetic activity [7]. The particular interest of these compounds resides, however, in their appurtenance to the series of local anesthetics the more so as it has been proposed [16, 17] that local anesthetics might block nerve conduction through attachment to axonal acetylcholine receptors. The study of the conformational relationship between these molecules and acetylcholine is therefore of obvious interest. Preliminary theoretical PCIO indications on the conformations to be expected for such molecules have been presented [18]. Here we probe deeper into this problem.

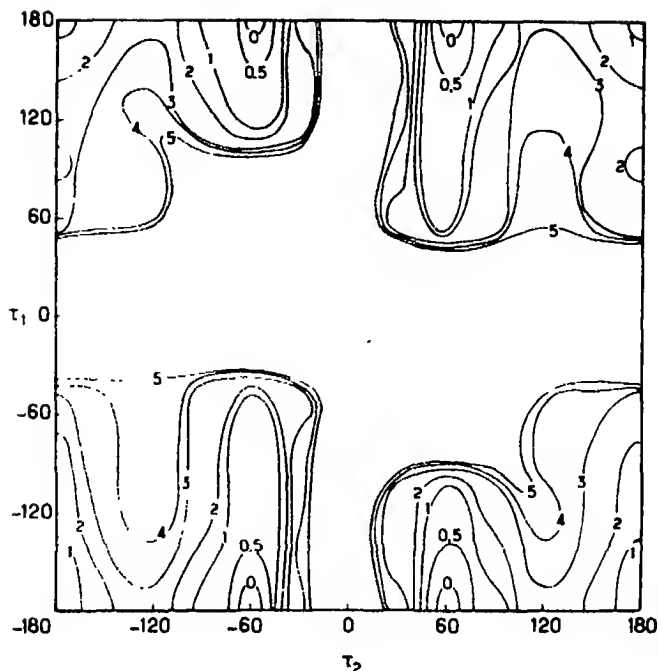


Fig. 8 Conformational energy map for compound VI. Isocnergy curves in kcal/mole with respect the global minimum taken as energy zero

Figures 8 and 9 present the conformational energy maps for the neutral VI and ionized VII forms, respectively, of the representative model of local anesthetics related to acetylcholine. (In building these maps and the following ones in this series the phenyl group was fixed coplanar with the plane of the ester group following the indications of the calculation of Ref. [18] and in agreement with the results of X-ray studies on procaine, *vide infra*). Figure 8 indicates a preference of the neutral form for the *gauche* conformation ( $\tau_1 = 180^\circ$ ,  $\tau_2 = \pm 60^\circ$ ) identical to the preferred form of acetylcholine with, however, a local energy minimum for the fully extended form ( $\tau_1 = \tau_2 = 180^\circ$ ) only 1 kcal/mole above the global minimum and a reduced energy barrier (2–3 kcal/mole) between the two. The ionized form (Fig. 9) manifests the same global minimum and shows also a local energy minimum for the extended form but which is now at 3 kcal/mole above the global one with a barrier of 4 kcal/mole between the two. From these results it may thus be inferred that the quaternization of the amino nitrogen increases the preference for the *gauche* conformation. It may be useful to remark here that although the identification of the active form in this class of drugs is still controversial, recent evidence seems to favor the cationic form as the active one (for references see [18]).

Recent experimental informations obtained with the use of the nuclear magnetic resonance technique [19, 20] confirm these theoretical conclusions: they indicate that while VI exists in solution 67% in the *gauche* forms and 33% in the *trans* form, (a situation which because of the possibility of two nearly equi-

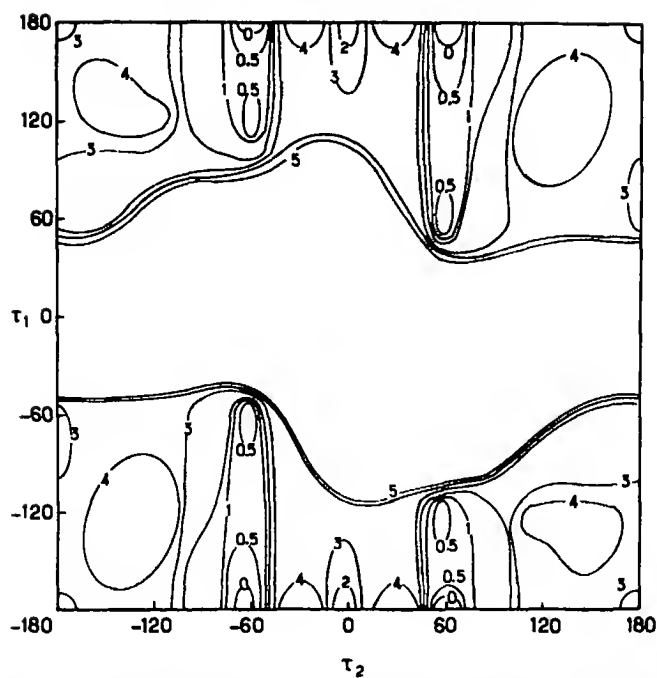


Fig. 9. Conformational energy map for compound VII. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

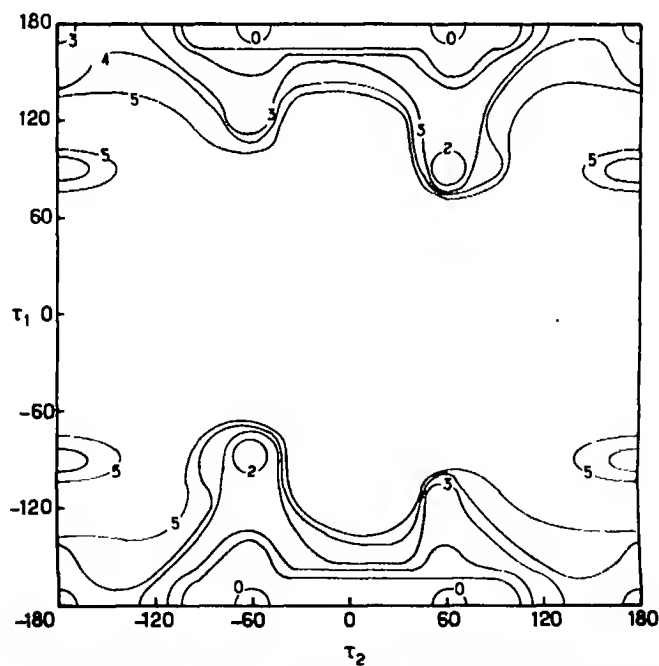


Fig. 10. Conformational energy map for compound IX. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

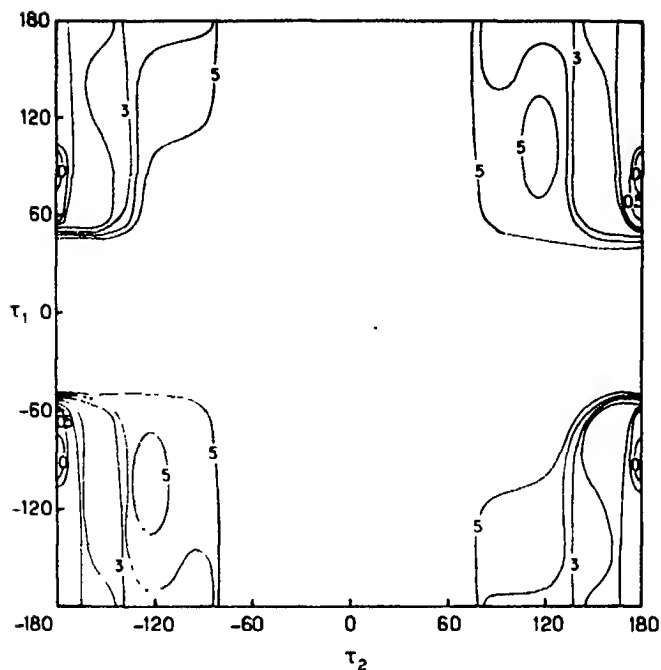


Fig. 11. Conformational energy map for compound X. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

valent *gauche* forms, indicates the near energetical equivalence of the three rotamers). VII exists in the same conditions exclusively in the *gauche* form.

The conformational relationship between derivatives of acetylcholine and those of local anesthetics in solution is still strengthened by the study of their sulfur analogs [21]. We have already discussed previously [22] the case of acetylthiocholine, VIII, which exists in the *trans* form both in the crystal [23] and in solution [24, 25] and whose PCIO conformational energy map is substantially different from that of acetylcholine, presenting a unique global energy minimum for the *trans* form. On the other hand, the replacement of the carbonyl oxygen by sulfur, leading to acetylthionecholine IX, yields the conformational map indicated in Fig. 10, which conserves the global minimum for the *gauche* conformation and which is, in fact, the conformation observed in solution [19]. The variations which are observed upon sulfur substitution in the series of the local anesthetics in solution parallel completely the variations observed in the acetylcholine series [21]. The corresponding conformational energy maps substantiate this analogy entirely. They are not reproduced here for the sake of saving space because of this very analogy. We may just add that when the two oxygens of such compounds are replaced by sulfur atoms, as in X, the conformational influence of the replacement of the acyloxy oxygen seems to be dominant (Fig. 11) and the compound is expected to exist largely in the *trans* form. It does so in 70% [18].



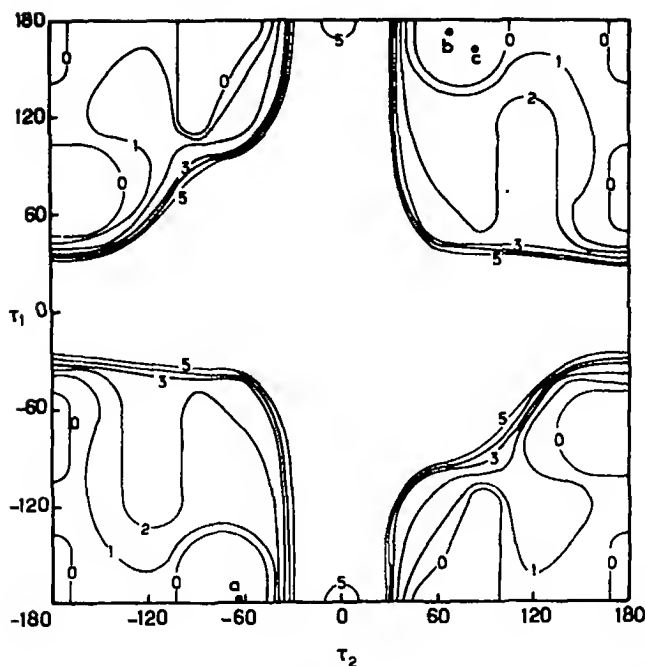


Fig. 12. Conformational energy map for procaine XI. Isoenergy curves in kcal/mole with respect to the global energy minimum taken as energy zero. ● Experimental conformations of procaine in crystals of: a 1:1 procaine-bis-*p*-nitrophenyl phosphate complex [27], b procaine hydrochloride [26, 28], c conformation of 2-diethylaminoethyl-*p*-methoxybenzoate hydrochloride [29]

Finally, in this series of compounds, we have also computed the conformational energy map for the effective local anesthetic procaine, XI, whose structure has been recently studied by X-ray crystallography by a number of authors [26-28]. Also studied by X-rays was the related 2-diethylamino-ethyl-*p*-methoxybenzoate hydrochloride [29]. Using as input data the crystallographic results for procaine hydrochloride [28], we obtain the conformational energy map of Fig. 12. The general aspect of the map (the contours of the stability zone within, say, the 3 kcal/mole isoenergy curve) is practically identical to that of acetylcholine. We observe, however, a degeneracy of the global energy minimum which is associated both with the *gauche* and *trans* forms. The minimum for the *gauche* form covers a somewhat larger area and may thus perhaps be considered as more probable. Whatever it be, the experimental conformations are all, as indicated in Fig. 12, of the *gauche* type and it is striking to observe that this major conformational feature of procaine is preserved in different crystal environments. It is also preserved in solution [30]. This situation indicates that intramolecular interactions responsible for these features compete successfully with intermolecular and environmental factors, as they, in fact, do frequently [13].

It may be interesting to add that the crystal data confirm also our assumptions, based on previous studies, on the planarity of the ester group and its coplanarity with the phenyl ring, coplanarity which is preserved in solution [31].

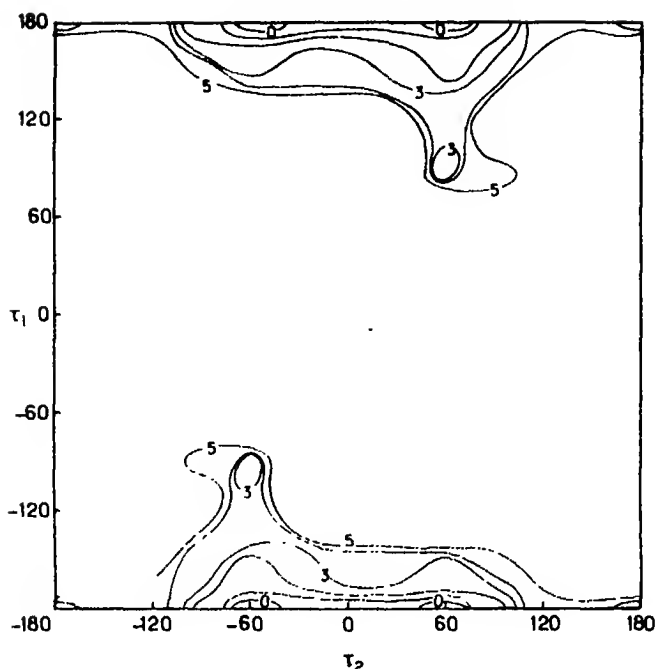


Fig 13. Conformational energy map for compound XII. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

There appears therefore to exist definite evidence, both theoretical and experimental, indicating that acetylcholine analogs and protonated local anesthetics are conformationally similar. This, by itself, does not mean, of course, that these similar conformational features are involved in the activities of these two groups of molecules, the less so as no proofs exists that these conformations, although characteristic of the structure in crystal and in solution, are also the ones involved in the interaction with the biological receptor(s). Truly, it was at a time proposed by some investigators [32, 33] that the *gauche* conformation is essential for the ability of cholinergic compounds to initiate a nerve impulse and similarly by others [16, 17] that the *gauche* conformation of local anesthetics may be an important feature necessary for the effector-receptor interaction and the blocking of the nerve impulse. Biological studies, however, conducted in particular by Mautner *et al.* [19–21] and by Partington *et al.* [34] show no consistent correlation between the conformations of molecules triggering or blocking conduction of the nerve impulse or affecting electrically excitable membranes and their potency. Conformational factors alone therefore probably do not determine neither the activity of cholinergic agonists nor that of local anesthetics. Electronic structure may be quite important in this field. In fact, it is frequently considered [6, 7] that the distribution of the electronic charges in particular in the carbonyl bond is important in the action of local anesthetics and that only slight complementarity is needed between these drugs and their receptor.

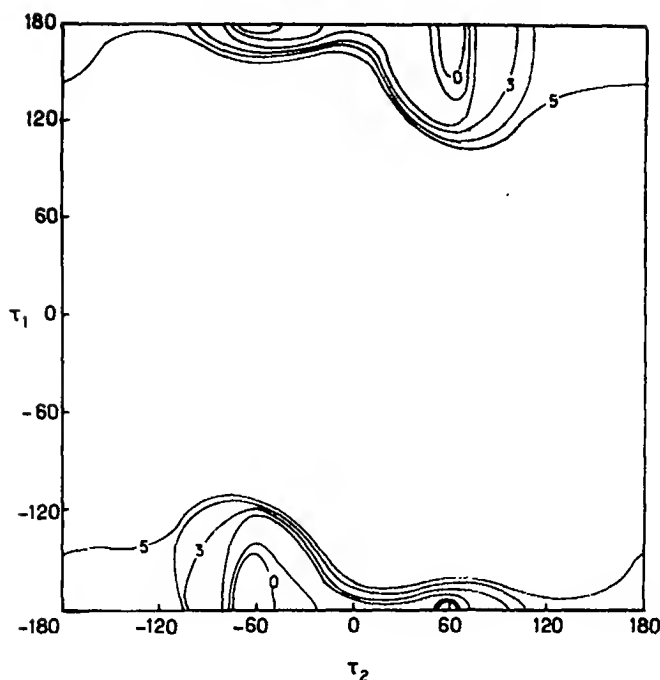


Fig. 14. Conformational energy map for compound XIII. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

### C. Choline Ethers

A more profound modification of the ester terminal of the acetylcholine skeleton is present in choline ethers and it appeared therefore interesting to investigate also the conformational characteristics of representative compounds of this series. Simple alkyl ethers manifest parasympathomimetic activity [9], the phenyl ether is a potent nicotine like ganglion stimulant, while its 2,6-dimethyl analogue has muscarine like activity [35].

We have constructed conformational energy maps for the last two molecules of this class, choline phenyl ether XII and choline-2,6 xylyl ether (xylocholine) XIII and for the simplest alkyl ether, the methyl ether XIV. In the calculations on XII, the phenyl group was assumed to be in the plane of the C—O—C atoms (following the indications of Ref. [36]), in the calculations on XIII to be perpendicular to that plane (following the indications of Ref. [35]).

The results are indicated in Figs. 13–15. They all show a strong preference for the *gauche* form analogous to that of acetylcholine. XIII and XIV are, however, devoid of a local minimum for the *trans* form. A very restricted such local minimum exists in XII, 4 kcal/mole above the global one. In XII and XIII the zone around  $\tau_1 = 180^\circ$ ,  $\tau_2 = 90^\circ - 180^\circ$  is about 1 kcal/mole higher than in acetylcholine. The conformationally stable zone of XII is, moreover, particularly restricted. The conformational properties of these molecules are thus similar to

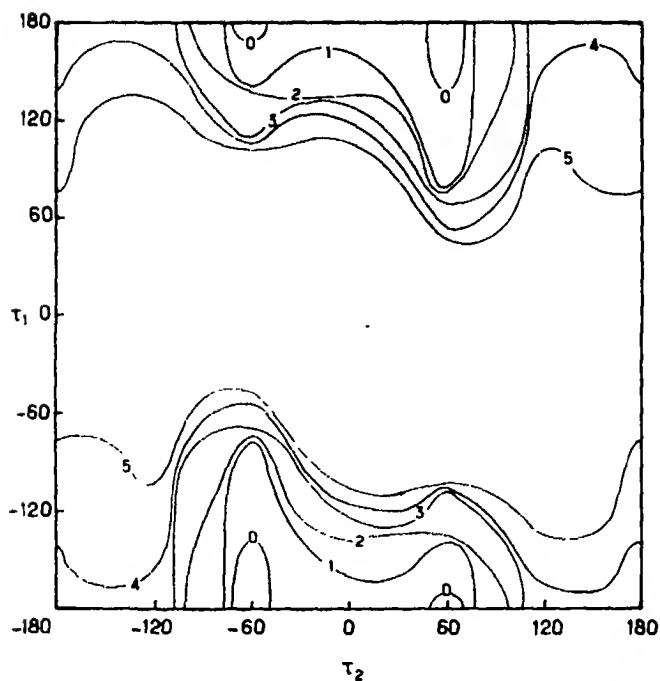
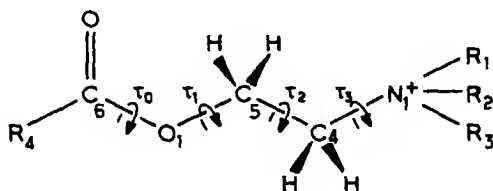


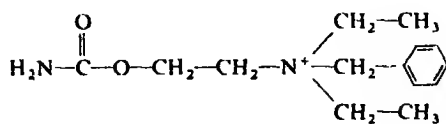
Fig. 15. Conformational energy map for compound XIV. Isoenergy curves in kcal/mole with respect to the global minimum taken as energy zero

those of acetylcholine with respect to the most stable conformer but may differ from it as to the possibility of other conformers.

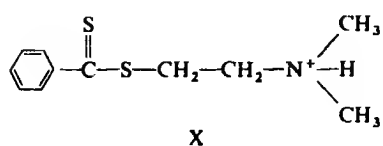
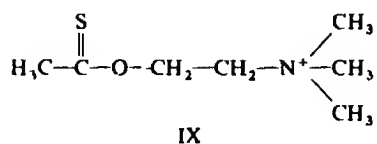
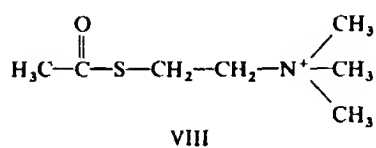
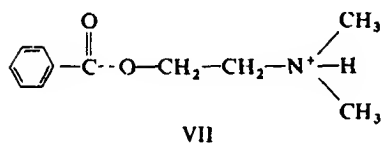
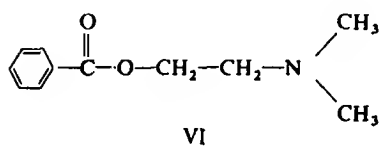
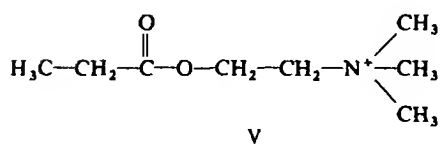
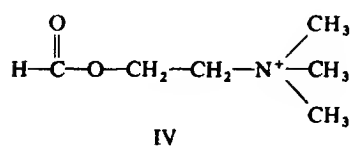
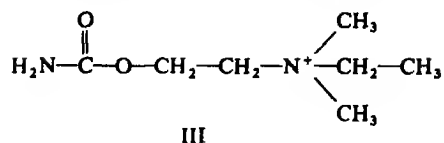
The available experimental data confirm the essentials of the theoretical results: NMR experiments show that the three ethers exist at about 80–90% in the *gauche* form in solution [34] and this is also the conformation observed for XI in the crystal [35].

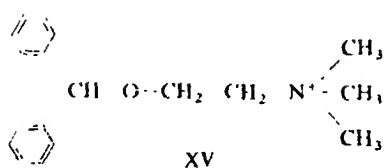
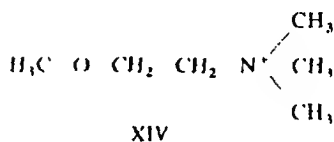
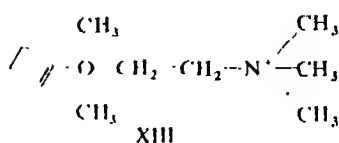
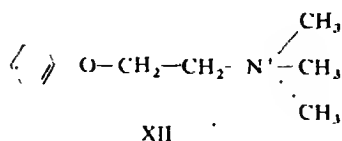
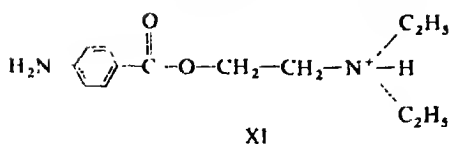


I



II





Finally the last compound investigated was the quaternary benzhydryl ether XV. Compounds of this class are competitive reversible antagonists at the muscarine receptor [37]. Detailed account about the conformational properties of this molecule will be given in a forthcoming paper on the structure of anti-histamic drugs. Suffice it to report here that the *gauche* and *trans* conformations are competitive in this compound.

It is thus manifest that conformational analogies or discrepancies alone cannot account in this series of molecules for the fine details of their biological activity.

### References

1. Pullman, B., Courrière, Ph., Coubeils, J. L.: *Molecular Pharmacology* 7, 391 (1971)
2. Pullman, B., Courrière, Ph.: *Molecular Pharmacology* 9, 612 (1972)
3. Pauling, P.: In: Bergmann, E. D., Pullman, B. (Eds.): *Conformation of biological molecules and polymers*. Proceedings of the 5<sup>th</sup> Jerusalem Symposium. New York: Academic Press 1973, p. 505

4. Shefter, E.: In: Triggie, D. J., Moran, J. F., Barnard, E. A. (Eds.): *Cholinergic ligand interactions*, p. 83. New York: Academic Press 1971
5. Kier, L. B.: *Molecular orbital theory in drug research*. New York: Academic Press 1971
6. Korolkovas, A.: *Essentials of molecular pharmacology*. New York: Wiley-Interscience 1970
7. Ariens, E. J., Simonis, A. M., Van Rossum, J. M.: In: Ariens, E. J. (Ed.): *Molecular pharmacology*, Vol. 1, p. 119. New York: Academic Press 1964
8. Herdktlotz, J. K., Sass, R. L.: *Biochem. Biophys. Res. Comm.* **40**, 583 (1970)
9. Ing, H. R.: *Science* **109**, 264 (1949)
10. Moran, J. F., Triggie, D. J.: In: Triggie, D. J., Moran, J. F., Barnard, E. A. (Eds.): *Cholinergic ligand interactions*, p. 119. New York: Academic Press 1971
11. Babeau, A., Barrans, Y.: *Compt. Rend. Acad. Sci. Paris* **270 C**, 609 (1970)
12. Barrans, Y.: Thesis, University of Bordeaux 1971
13. Pullman, B., Courrière, Ph.: In: Bergmann, E. D., Pullman, B. (Eds.): *Conformation of biological molecules and polymers. Proceedings of the 5<sup>th</sup> Jerusalem Symposium*. New York: Academic Press 1973, p. 547
14. Barrans, Y., Clastre, J.: *Compt. Rend. Acad. Sci. Paris* **270**, 306 (1970)
15. Conti, F., Damiani, A., Pietronero, C., Russo, N.: *Nature New Biology* **233**, 232 (1971)
16. Bartels, E., Nachmansohn, D.: *Biochem. Ztschr.* **342**, 359 (1965)
17. Bartels, E.: *Biochim. Biophys. Acta* **109**, 194 (1965)
18. Coubeils, J. L., Pullman, B.: *Molecular Pharmacology* **8**, 278 (1972)
19. Mautner, H. G., Dexter, D. D., Low, B. W.: *Nature New Biology* **238**, 87 (1972)
20. Makriyannis, A., Sullivan, R. F., Mautner, H. G.: *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3416 (1972)
21. Chu, S. H., Hillman, G. R., Mautner, H. G.: *J. Med. Chem.* **15**, 760 (1972)
22. Pullman, B., Courrière, Ph.: *Molecular Pharmacology* **8**, 371 (1972)
23. Shefter, E., Mautner, H. G.: *Proc. Natl. Acad. Sci. U.S.A.* **63**, 1253 (1969)
24. Culvencor, C. C. J., Ham, N. S.: *Chem. Comm.* **242** (1970)
25. Cushley, R. J., Mautner, H. G.: *Tetrahedron* **26**, 2151 (1970)
26. Beall, R., Herdktlotz, J., Sass, R. L.: *Biochem. Biophys. Res. Comm.* **39**, 329 (1970)
27. Pletcher, M. S. J., Gustaffson, B.: *Acta Cryst.* **B26**, 114 (1970)
28. Dexter, D. D.: *Acta Cryst.* **B28**, 77 (1972)
29. Beall, R., Sass, R. L.: *Biochem. Biophys. Res. Comm.* **40**, 833 (1970)
30. Chu, S. H., Hillman, G. R., Mautner, H. G.: *J. Med. Chem.* **15**, 760 (1972)
31. Gansow, O. A., Beckenbaugh, W. M., Sass, R. L.: *Tetrahedron* **28**, 2691 (1972)
32. Sundaralingam, M.: *Nature* **217**, 35 (1968)
33. Baker, R. W., Chothia, C. H., Pauling, P., Petcher, I. J.: *Nature* **230**, 439 (1971)
34. Partington, P., Feeney, J., Burgen, A. S. V.: *Molecular Pharmacology* **8**, 269 (1972)
35. Coggon, P., McPhail, A. T., Roc, A. M.: *Nature* **224**, 1200 (1969)
36. Coubeils, J. L., Courrière, Ph., Pullman, B.: *J. Med. Chem.* **15**, 453 (1972)
37. Moran, J. F., Triggie, D. J.: In: Danielli, J. F., Moran, J. F., Triggie, D. J. (Eds.): *Fundamental concepts in drug-receptor interactions*, p. 123. New York: Academic Press 1970

Prof. Dr. B. Pullman  
Institut de Biologie Physico-Chimique  
Laboratoire de Biochimie Théorique  
associé au C.N.R.S.  
13, rue Pierre et Marie Curie  
F-75 Paris 5<sup>e</sup>, France





## *Ab initio* Calculations on Small Hydrides Including Electron Correlation

### IX. Equilibrium Geometries and Harmonic Force Constants of HF, OH<sup>-</sup>, H<sub>2</sub>F<sup>+</sup> and H<sub>2</sub>O and Proton Affinities of F<sup>-</sup>, OH<sup>-</sup>, HF and H<sub>2</sub>O

Hans Lischka\*

Institut für Physikalische Chemie und Elektrochemie, Abteilung für Theoretische Chemie, Universität  
Karlsruhe, Kaiserstraße 12, D-7500 Karlsruhe, Germany

Received March 1, 1973/May 11, 1973

Near Hartree-Fock *ab initio* SCF and correlation energy calculations based on the IEPA-PNO approximation are reported for F<sup>-</sup>, HF, OH<sup>-</sup>, H<sub>2</sub>O, H<sub>2</sub>F<sup>+</sup> and H<sub>3</sub>O<sup>+</sup> using GTO-basis sets. The SCF values for the equilibrium geometries and harmonic force constants are corrected in the desired direction by inclusion of the correlation energy. The SCF errors are, however, always overcompensated. The agreement with experiment is improved for the symmetric stretching force constants of HF and H<sub>2</sub>O, but bond lengths are nearly the same amount too long in the IEPA approximation as they are too short in the SCF calculations. In addition thereto protonation energies are computed and compared with experimental measurements.

**Key words:** Force constants – Proton affinities – Hydrides – Electron correlation.

#### 1. Introduction

In previous papers of this series the effect of electron correlation on the molecular structure has been investigated for the hydrides of Li, Be, B and C. In the present paper computations on H<sub>2</sub>O, HF and their protonated and deprotonated ions are reported.

Although these compounds are well known in chemistry, accurate spectroscopic gas phase data exist only for F<sup>-</sup>, HF and H<sub>2</sub>O [1]. In the last years mass-spectrometric experiments involving OH<sup>-</sup> and H<sub>3</sub>O<sup>+</sup> and their higher solvation complexes have been performed [2], which give important supplementary information on these molecules that are well-known from studies in solution and in solid state. H<sub>2</sub>F<sup>+</sup> which is known to protonate even weak bases [3] has only been reported to exist as solute in the so-called "magic acid" [4].

Experimental information on the structure of atoms and molecules depends on the environment of the systems which are measured. In order to facilitate com-

\* On leave from: Institut für Theoretische Chemie, Universität Wien, Währingerstraße 38, A-1090 Wien, Austria.

parison of different experimental results, data for the isolated species are desirable. Theory gives this information although the degree of accuracy depends on the various approaches used. Our intention was to perform as accurate calculations as possible with reasonable computation time. We checked the accuracy of our results a) by comparison with more sophisticated calculations which are only possible for smaller systems and b) by comparison with existing experimental data for the molecules in the gas phase. Proceeding this way it is possible to give a consistent theoretical treatment of a whole series of molecules and to estimate the reliability of our conclusions.

## 2. Method of Computation

As in the other papers of this series the method used is based on the independent-electron-pair approximation (IEPA) [5, 6] and direct calculation of approximate natural orbitals of electron-pair functions developed by Kutzelnigg, Ahlrichs, and co-workers [7, 8].

After performing a conventional Hartree-Fock calculation the canonical orbitals are transformed to localized ones using the criterion of Boys [9]. For each doubly occupied localized orbital  $\varphi_R$  one calculates a pair function in its natural expansion form in the effective field of the other electrons together with the intrapair correlation energy contribution  $\epsilon_{RR}$ . Similarly, we find for pairs of different localized orbitals  $\varphi_R$  and  $\varphi_S$  singlet and triplet coupled pair functions in the effective field of the remaining electrons which give the interpair contributions  $\epsilon_{RS}$  and  $\epsilon_{RS}^1$ . The sum of the different intra- and interpair correlation energies is regarded as an approximation to the total correlation energy.

## 3. Choice of the Basis Set

For the O and F atom the (11s, 7p) GTO-basis set of Huzinaga [10] was used and contracted to a [7s, 4p] set, with the contraction coefficients taken from the atom. For the H atom we contracted Huzinaga's (6s) basis [11] to four groups<sup>2</sup>. Polarization functions of *d*-type on O and F and of *p*-type on H with optimized exponents were added. We simulate *p*- and *d*-functions by appropriate linear combinations of lobes which are shifted from the origin [12, 13]. Two basis sets were considered: for basis set *A* one set of *d*-functions (exponent 1.5) on F and O, respectively, and one set of *p*-functions (exp. 0.75) on H is used, for basis set *B* two *d*'s (exps. 0.75 and 3.0) and two *p*'s (exps. 0.65 and 2.6). Most of the calculations were done with basis set *A*, giving sufficient accuracy for a discussion of SCF- and IEPA effects on equilibrium geometries and force constants; only few selected points were computed with the larger basis *B*.

<sup>1</sup> In the following text we shall tabulate the sum  $\epsilon_{RS} = \epsilon_{RS} + \epsilon_{RS}^1$  only.

<sup>2</sup> In constructing the contracted functions always those with the largest exponents were grouped together, i.e. (5 1 1 1 1 1) for the (11s) set, on O and F, respectively, (4 1 1 1) for the (7p) set and (3 1 1 1) for (6s) set on H.

## 4. Results and Discussion

a) Equilibrium Geometries and Force Constants for HF, OH<sup>-</sup>, H<sub>2</sub>O and H<sub>2</sub>F<sup>+</sup>

The SCF energy and the pair correlation increments for the valence shell were calculated with basis *A* for HF and OH<sup>-</sup> at seven internuclear distances. The results for HF are collected in Table 1. If not stated otherwise, all energies are given in atomic units and are negative; the labels *b* and *n*<sup>3</sup> refer to bond pair and lone pair, respectively. The dependence of the correlation energy on the distance *r* is dominated by  $\epsilon_{bb}$  and  $\epsilon_{bn}$  which both increase in absolute value,  $|\epsilon_{nn}|$  is slightly increasing,  $|\epsilon_{nn'}|$  decreasing; an analogous situation is found for OH<sup>-</sup>. The correlation of the inner shell electrons have been shown to remain approximately constant in other calculations [14, 15] and should therefore not change the shape of the potential energy curve significantly. Therefore it is not taken into account here.

For H<sub>2</sub>O and H<sub>2</sub>F<sup>+</sup> the dependence of  $\epsilon_{nn}$ ,  $\epsilon_{bn}$  and  $\epsilon_{nn'}$  on *r* (keeping  $\alpha = \text{const}$ ) is the same as for HF and OH<sup>-</sup>. Variation of the bond angle  $\alpha$  (at *r* = const) shows the expected dominating increase of  $|\epsilon_{bb}|$  with decreasing  $\alpha$  (Table 2).

The Boys localization procedure yielded in all cases equivalent lone pair orbitals.

Table 1. Variation of SCF-, pair-correlation- and total energies (a.u.) with the bond distance *r* (a.u.) for HF

<i>r</i>	1.50	1.65	1.69	1.733	1.79	1.85	2.10
$E_{\text{SCF}}$	100.0457	100.0640	100.0650	100.0647	100.0624	100.0583	100.0287
$\epsilon_{bb}$	0.0257	0.0279	0.0285	0.0292	0.0301	0.0310	0.0350
$\epsilon_{nn}$	0.0210	0.0211	0.0211	0.0211	0.0212	0.0212	0.0213
$\epsilon_{bn}$	0.0266	0.0277	0.0280	0.0283	0.0287	0.0291	0.0305
$\epsilon_{nn'}$	0.0280	0.0279	0.0279	0.0279	0.0279	0.0279	0.0277
$\sum_{R < S} \epsilon_{RS}$	0.2523	0.2581	0.2595	0.2612	0.2632	0.2653	0.2735
$E_{\text{tot}}$	100.2980	100.3221	100.3245	100.3258	100.3256	100.3236	100.3022

Table 2. Variation of SCF-, pair-correlation- and total energies (a.u.) with the bond distance *r*<sub>OH</sub> (a.u.) and bond angle  $\alpha$  for H<sub>2</sub>O

	$\alpha = 104.5^\circ$	$r = 1.809$		$r = 1.809$	$\alpha = 90^\circ$	$120^\circ$
	$r = 1.70$	1.809	1.95			
$E_{\text{SCF}}$	76.0552	76.0593	76.0452	76.0522	76.0546	
$\epsilon_{bb}$	0.0280	0.0295	0.0316	0.0296	0.0291	
$\epsilon_{nn}$	0.0227	0.0228	0.0230	0.0229	0.0227	
$\epsilon_{bn}$	0.0230	0.0244	0.0263	0.0259	0.0235	
$\epsilon_{nn'}$	0.0260	0.0265	0.0272	0.0268	0.0265	
$\epsilon_{nn''}$	0.0290	0.0288	0.0285	0.0281	0.0295	
$\sum_{R < S} \epsilon_{RS}$	0.2572	0.2640	0.2729	0.2661	0.2629	
$E_{\text{tot}}$	76.3124	76.3232	76.3180	76.3183	76.3176	

Received November 1979

38969

The net effect of correlation is to increase the bond length (the same behaviour was observed in other cases [15]) and to decrease the bond angle.

The effects of correlation on bond distances and harmonic force constants are clearly demonstrated in Tables 3 and 4. For HF and OH<sup>-</sup> we fitted a function of the form  $\sum_i a_i x^i e^{-\alpha_i x}$  ( $\alpha_i$  also optimized) to 7 points at different internuclear distances and computed the equilibrium geometries and the force constants analytically. In the case of the triatomic molecules H<sub>2</sub>O and H<sub>2</sub>F<sup>+</sup> 17–20 points on the two-dimensional energy surface were calculated to determine the force constants of the totally symmetric vibrations; in analogy to the linear case these points were interpolated by the following sum:  $\sum_{ij} a_{ij} x^i y^j e^{-\alpha_i x} e^{-\beta_j y}$ . In addition to that we considered for H<sub>2</sub>O also the antisymmetric vibration.

For the triatomic molecules two equivalent forms of the harmonic potential are used [1c, part II]:

$$2V = 2V_0 + F_{11}(r_1^2 + r_2^2) + F_{33}\alpha^2 + 2F_{12}r_1r_2 + 2F_{13}(r_1 + r_2)\alpha$$

and

$$2V = 2V_0 + c_{11}S_1^2 + 2c_{12}S_1S_2 + c_{22}S_2^2 + c_{33}S_3^2.$$

The symmetry coordinates  $S_1$ ,  $S_2$  and  $S_3$  are given in terms of the internal displacement coordinates  $r_1$ ,  $r_2$  and  $\alpha$ :

$$S_1 = \alpha$$

$$S_2 = \frac{1}{\sqrt{2}}(r_1 + r_2)$$

$$S_3 = \frac{1}{\sqrt{2}}(r_1 - r_2).$$

The  $c$ 's and  $F$ 's are related in the following way:

$$F_{11} = -\frac{c_{22} + c_{33}}{2}$$

$$F_{12} = \frac{c_{22} - c_{33}}{2}$$

$$F_{13} = \frac{c_{12}}{\sqrt{2}}$$

and

$$F_{33} = c_{11}.$$

A comparison of the SCF results (see Table 3 and 4) with experimental values shows the following characteristics: bond lengths are too short by 2% and stretching force constants too large by about 15%. Inclusion of the electron correlation by means of the IEPA-PNO approximation shifts the calculated equilibrium geometries and force constants in the desired direction but always overcorrects

Table 3. Calculated and experimental bond distances  $R_0$ (Å) and harmonic force constants  $k_0$ (mdyn/Å) for HF and OH<sup>-</sup>

		Hartree-Fock limit [16, 17]	This work		exp <sup>a</sup>
			SCF	IEPA	
HF	$R_0$	0.897	0.900	0.929	0.917
	$k_0$	11.261	11.36	9.407	9.657
OH <sup>-</sup>	$R_0$	0.942	0.945	0.974	0.970
	$k_0$	9.334	9.001	7.438	—

<sup>a</sup> experimental values were taken from [16] and [17].Table 4. Equilibrium geometries, harmonic force constants (mdyn/Å) and harmonic vibration frequencies (cm<sup>-1</sup>) for H<sub>2</sub>O and H<sub>2</sub>F<sup>+</sup><sup>a</sup>

		Dunning <i>et al.</i> [19]	This work		exp	Ref.
		SCF	SCF	IEPA		
H <sub>2</sub> O	$R_0$ (Å)	0.941	0.942	0.973	0.958	[21]
	$\alpha_0$	106.6	106.5	103.3	104.5	
	$F_{11}$	9.38	9.31	8.37	8.45	
	$F_{12}$	-0.078	0.395	-0.550	-0.101	[22]
	$F_{13}/R_0$	0.249	0.232	0.222	0.227	
	$F_{33}/R_0^2$	0.867	0.821	0.722	0.761	
	$\bar{\nu}_1$	—	1695.1	1590.3	1653.9	[23]
	$\bar{\nu}_2$ } $A_1$	—	4143.4	3725.5	3825.3	
	$\bar{\nu}_3$ } $B_1$	—	4045.3	4034.5	3935.6	
	$\bar{\nu}_2$ } $A_1$	—	3790.8	3475.4	—	
H <sub>2</sub> F <sup>+</sup>	$R_0$ (Å)	—	0.942	0.970	—	
	$\alpha_0$	—	116.0	111.8	—	
	$c_{11}/R_0^2$	—	0.549	0.543	—	
	$c_{22}$	—	8.227	6.896	—	
	$c_{12}/R_0$	—	0.0857	0.127	—	
	$\bar{\nu}_1$	—	1384.2	1378.2	—	
	$\bar{\nu}_2$ } $A_1$	—	3790.8	3475.4	—	

<sup>a</sup> For H<sub>2</sub>F<sup>+</sup> we did not consider the antisymmetric vibration — therefore only the force constants corresponding to the symmetry coordinates are given.

the SCF errors. The agreement with experiment is improved for the symmetric stretching force constants (deviation from experimental value approx. -5%); the bond distances (with IEPA), however, are nearly the same amount too long as they are too short in the SCF calculations (see Tables 3 and 4). Recently Meyer [18] showed also for the CH<sub>4</sub> molecule by a comparison of his PNO-CI and CEPA methods with the IEPA approach that IEPA tends to overestimate the correlation energy changes with increasing bond length.

In the case of H<sub>2</sub>O (Table 4) our SCF force constants agree with other treatments [19, 20], with the exception of  $F_{12}$  for which we obtain the opposite sign.  $F_{12}$  is here calculated as the difference of two large numbers, an error of a few

percent in  $c_{22}$  and  $c_{33}$  is sufficient to change the sign of  $F_{12}$ . Here it is interesting to look at the effect of correlation in more detail: the constant  $c_{33}$  for the antisymmetric vibration is hardly influenced by correlation, because the stretching of one bond implies the shortening of the other, which causes the changes of the different pair energies to cancel each other approximately. In our calculations the constant  $c_{22}$  for the symmetric stretching vibration (and therefore also  $F_{11}$ ) decreases with inclusion of correlation, as one would expect from the results for diatomic hydrides (Table 3). Quite the same behaviour has also been observed by Staemmler and Jungen [24]<sup>4</sup> for the  $\text{BH}_2$  radial. We disagree with McLaughlin, Bender and Schaefer [25] who find no change in the stretching force constant  $F_{11}$  when correlation is included. In our SCF calculation the vibration with the largest frequency corresponds to symmetry  $A_1$ , which is in contradiction to the experimental result - correlation produces then the correct order of the frequencies although the overshooting of the IEPA correction is manifested here also.

Using the same size of the basis set for  $\text{H}_2\text{F}^+$  no new aspects appear which would make the computed results less reliable than those for  $\text{H}_2\text{O}$ . To our knowledge no experimental values exist for the equilibrium geometry and harmonic force constants; a direct comparison of the calculated harmonic frequencies with the IR-spectrum measured by Hyman *et al.* [4] is not possible because of solvation effects and inharmonic contributions to the vibrations which were not considered in our calculations.

### b) Energies of Protonation

At the equilibrium geometries SCF and IEPA calculations were performed with the larger basis set *B*. Results are given in Table 5. Our SCF values for  $\text{F}^+$ ,  $\text{HF}$  and  $\text{OH}^+$  are all less than 0.005 a.u. above the Hartree-Fock limit, 0.006 a.u. for  $\text{H}_2\text{O}$  (estimated Hartree-Fock limit by Clementi and Popkie [28] - 76.068 a.u.), the same should hold for  $\text{H}_2\text{F}^+$ . It is difficult to compare our correlation energies with values of the other authors (Table 5) because they used canonical orbitals for their CI calculation in contrast to our localized ones. It is known that the pair energies are not invariant under unitary transformations of the SCF-MO's [32]. Moreover our basis set does not contain *f*-functions.

For  $\text{H}_3\text{O}^+$  a pyramidal configuration is found to be slightly more stable than the planar one. The height of the barrier of inversion is too small (0.0015 a.u. in SCF) to allow a definite decision whether it exists at all, inclusion of correlation enlarges this barrier [33]. This uncertainty, however, does not affect our calculated proton affinities (a detailed study of the structure of  $\text{H}_3\text{O}^+$  is in preparation [33]).

The proton affinity is defined here as the difference between the energies of the protonated species and the parent molecule at the minimum of the potential energy surface, i.e. we have not included the zero point energy. Protonation energies from SCF calculations with various smaller GTO-basis sets than ours

<sup>4</sup> There is a misprint in Table 3 of [24].  $F_{12}$  should read as -0.14 instead of 0.14 (private communication).

Table 5. Comparison of different SCF and CI results with experimental values

	$r(\text{a.u.})$	$\alpha$	IEPA		$E_{\text{corr}}^{\text{K}}$ (K-shell)	$E_{\text{corr}}^{\text{V}}$ (valence shell)	$E_{\text{corr}}^{\text{T}}$ (total)	Variational CI		exp energy <sup>b</sup>
			$E_{\text{SCF}}$	$E_{\text{corr}}^{\text{K}}$				$E_{\text{corr}}$	$E_{\text{tot}}$	
F <sup>-</sup>										
This work			99.4571 <sup>c</sup>	0.0495	0.3007	0.3502	99.8073	—	—	—
Weiss [26]			99.4594	0.0598	0.3219	0.3817	99.8411	—	—	99.858
HF										
This work	1.733	—	100.0660	0.0485	0.2860	0.3345	100.4005	—	—	—
Cade and Huo [16]	1.733	—	100.0703	—	—	—	—	—	—	100.477
Bender and Davidson [27]	1.733	—	100.0703	0.0563	0.3083	0.3646	100.4349	0.2861	100.3564	—
OH <sup>-</sup>										
This work	1.781	—	75.4128	0.0492	0.3009	0.3501	75.7629	—	—	—
Cade [17]	1.781	—	75.4175	—	—	—	—	—	—	75.781
H <sub>2</sub> O										
This work	1.809	104.5	76.0620	0.0479	0.2832	0.3311	76.3931	—	—	—
Clementi [28]	1.809	104.5	76.0659	—	—	—	—	—	—	76.436
Meyer [29]	1.809	104.5	76.0628	0.0560	0.3020	0.3580	76.4208	0.3056	76.3683	—
H <sub>2</sub> F <sup>+</sup>										
This work	1.81	113	100.2612	—	0.2833	0.2833	100.5444	—	—	—
Hopkinson <i>et al.</i> [30]	1.733	105	100.2082	—	—	—	—	—	—	—
H <sub>3</sub> O <sup>+</sup>										
Basis set B	1.81	10	76.3418	—	—	—	—	—	—	—
Basis set A	1.85	15	76.3386	—	0.2629	0.2629	76.6015	—	—	—
Dierksen [31]	1.82	0	76.3280	—	—	—	—	—	—	—

<sup>a</sup> Sum of the *intra* and inter pair contributions of the inner (K-) shell to the correlation energy. In this case all of the 11s and 7p functions were decontracted.<sup>b</sup> The values are taken from table 3 of [30] and are defined as the sum of the atomic-, dissociation- and zero-point vibrational energy, but corrected for relativistic effects.<sup>c</sup> One set of *p*-functions with exponents 0.08 was added.

Table 6. Calculated and experimental proton affinities (in kcal/mole) (all energies are negative)

	This work		With correlation <sup>d</sup>		Hartree-Fock limit <sup>a</sup>	Hopkinson <i>et al.</i> [30]	exp <sup>b</sup>		
	SCF								
	Basis set		Basis set						
	<i>A</i>	<i>B</i>	<i>A</i>	<i>B</i>					
F	380.9	382.1	374.3	372.9	383.5	386.6	388.6 <sup>c</sup>	361	368
HF	121.8	122.5	121.8	120.8	-	108.4	---	---	---
OH	406.4	407.5	398.2	396.4	408.0	403.2	411.1 <sup>c</sup>	370	392
H <sub>2</sub> O	176.6	175.6	174.0	-	-	181.2	---	151	181

<sup>a</sup> The energy for F is taken from [35], for HF and OH from [16] and [17] and for H<sub>2</sub>O from [28].

<sup>b</sup> Values taken from Table 5 of [30].

<sup>c</sup> From spectroscopic data (see also Table 5, last column).

<sup>d</sup> Only the correlation energy for the valence shell is taken into account.

are compared by Hopkinson *et al.* [30]. Schuster [34] gives semiempirical (CNDO/2) values. The SCF-approximation gives already the main contributions, inclusion of correlation modifies these results only to a very small extent (Table 6). Even over the whole series of molecules given in Table 5 correlation does not change enough to have a significant influence on energy differences in the order of magnitude of some 100 kcal/mole.

Our basis sets are supposed to have enough flexibility to insure an approximately equal accuracy of the total energy for the isoelectronic series of systems considered here, as far as protonation energies are concerned. This means that our values for the proton affinities should be of comparable quality for the whole series. Experimental results, especially for H<sub>2</sub>O vary over a relative wide range depending on the method of determination; from our point of view the larger values are favoured for H<sub>2</sub>O (this is also found in recent measurements of De Paz *et al.* [2b]).

## 5. Conclusions

The present investigations show what improvements can be obtained for the calculation of equilibrium geometries and force constants by taking into account electron correlation in the framework of the IEPA-PNO approximation, but also what limits are connected with this method. The best results are obtained for the symmetric stretching force constants – the too large SCF values (+ 15%) are decreased substantially and then agree reasonably with experiment (– 5%). The bond distances, however, are approximately the same amount too long with inclusion of the IEPA correlation energy as they are too short in the SCF calculation. This overcorrection of the SCF error is not so large for the bond



angle of  $\text{H}_2\text{O}$ . The comparison of our calculated results with experimental values indicates also the reliability of our predictions for the molecular properties of the experimentally less known systems  $\text{H}_2\text{F}^+$  and  $\text{H}_3\text{O}^+$ .

**Acknowledgements.** The author is indebted to Prof. Dr. W. Kutzelnigg and his group for the kind hospitality and numerous stimulating discussions.

The numerical computations were performed at the „Rechenzentrum der Universität Karlsruhe“; the program for the IEPA calculation was written by Dr. Ahlrichs, Dr. Staemmler and Dr. Jungen, that for the analytical fitting of the potential curves by Dr. Ahlrichs.

A research fellowship granted by the „Alexander von Humboldt-Stiftung“ is gratefully acknowledged.

This research was supported in part by the „Deutsche Forschungsgemeinschaft“ and the „Fonds der Chemie“.

## References

- 1a. Moore, C. E.: Atomic energy levels, National Bureau of Standards, Circ. No. 467 Vol I (1969)
- 1b. Berry, R. S., Reimann, C. W.: J. Chem. Phys. **38**, 1540 (1963)
- 1c. Herzberg, G.: Molecular spectra and molecular structure, I, II, III. D. van Nostrand Comp. Inc. 1967, 1962, 1966
- 2a. Good, A., Durden, D. A., Kabarle, P.: J. Chem. Phys. **52**, 212, 222 (1970)
- 2b. De Paz, M., Leventhal, J. J., Friedman, L.: J. Chem. Phys. **51**, 3748 (1969)
- 2c. De Paz, M., Guidoni Giardini, A., Friedman, L.: J. Chem. Phys. **52**, 687 (1970)
3. Olah, G. A., Klopman, G., Schlosberg, R. H.: J. Am. Chem. Soc. **91**, 3261 (1969)
- 4a. Hyman, H. H., Quarterman, L. A., Kilpatrick, M., Katz, J. J.: J. Phys. Chem. **65**, 123 (1961)
- 4b. Gillespie, R. J., Moss, K. C.: J. Chem. Soc. (London) A, **1966**, 1170
5. Nesbet, R. K.: Phys. Rev. **155**, 51 (1967)
6. Sinanoglu, O.: J. Chem. Phys. **36**, 706 (1962); **36**, 3198 (1962); – Advan. Chem. Phys. **6**, 315 (1964)
7. Kutzelnigg, W.: Theoret. Chim. Acta (Berl.) **1**, 327 (1963)
8. Jungen, M., Ahlrichs, R.: Theoret. Chim. Acta (Berl.) **17**, 339 (1970)
9. Boys, S. F. in: Quantum theory of atoms, molecules and the solid state, P. O. Löwdin (Ed.), p. 253. New York: Interscience 1967
10. Huzinaga, S., Sakai, Y.: J. Chem. Phys. **50**, 1371 (1969)
11. Huzinaga, S.: J. Chem. Phys. **42**, 1293 (1965)
12. Preuss, H. W.: Z. Naturforsch. **11a**, 823 (1965)
13. Whitten, J. L.: J. Chem. Phys. **44**, 359 (1966)
14. Pipano, A., Gilman, R. R., Shavitt, I.: Chem. Phys. Lett. **5**, 285 (1970)
15. e.g. Das, G., Wahl, A. C.: J. Chem. Phys. **44**, 87 (1966). – Gélus, M., Ahlrichs, R., Staemmler, V., Kutzelnigg, W.: Theoret. Chim. Acta (Berl.) **21**, 63 (1971)
16. Cade, P. E., Huo, W. H.: J. Chem. Phys. **47**, 614 (1967)
17. Cade, P. E.: J. Chem. Phys. **47**, 2390 (1967)
18. Meyer, W.: J. Chem. Phys. **58**, 1017 (1973)
19. Dunning, Th., Pitzer, P. M., Aung, S.: J. Chem. Phys. **57**, 5044 (1972)
20. Pulay, P.: Mol. Phys. **21**, 329 (1971)
21. Sutton, L. E. (Ed.): Table of interatomic distances and configuration in molecules and ions, London: The Chemical Society, Burlington House, W.1.1958
22. Kuchitsu, K., Morino, V.: Bull. Chem. Soc. Japan **38**, 814 (1965)
23. Darling, B. T., Dennison, D. M.: Phys. Rev. **57**, 128 (1940)
24. Staemmler, V., Jungen, M.: Chem. Phys. Letters **16**, 187 (1972)
5. McLaughlin, D. R., Bender, Ch., Schaefer III, H. F.: Theoret. Chim. Acta (Berl.) **25**, 352 (1972)
6. Weiss, A. W.: Phys. Rev. A **3**, 126 (1971)
7. Bender, Ch. F., Davidson, E. R.: Phys. Rev. A **183**, 23 (1969)
8. Clementi, E., Popkie, H.: J. Chem. Phys. **57**, 4870 (1972)
9. Meyer, W.: Intern. J. Quant. Chem. **5S**, 296 (1971)
0. Hopkinson, A. C., Holbrook, N. K., Yates, K., Csizmadia, I. G.: J. Chem. Phys. **49**, 3596 (1968)
1. Kraemer, W. P., Dierksen, G. H. F.: Chem. Phys. Letters **5**, 463 (1970)

32. Bender, C. F., Davidson, E. R.: Chem. Phys. Letters **3**, 33 (1970)
33. Dyczmons, V., Lischka, H.: to be published
34. Schuster, P.: Theoret. Chim. Acta (Berl.) **19**, 212 (1970)
35. Clementi, E.: Tables of atomic functions, supplement to IBM Journal of Research and Development **9**, 2 (1965)

Dr. H. Lischka  
Institut für Physikalische  
Chemie und Elektrochemie  
Abt. für Theoretische Chemie  
Universität Karlsruhe  
D-7500 Karlsruhe  
Kaiserstr. 12  
Federal Republic of Germany

## *Ab initio* Calculations on Small Hydrides Including Electron Correlation

### X. Triplet-Singlet Energy Separation and Other Properties of the $\text{CH}_2$ Radical

Volker Staemmler

Institut für Physikalische Chemie und Elektrochemie der Universität Karlsruhe,  
Karlsruhe, Kaiserstraße 12, Germany

Received April 3, 1973

The two lowest electronic states ( $^3B_1$  and  $^1A_1$ ) of the methylene radical ( $\text{CH}_2$ ) are calculated both in SCF-approximation and with the IEPA-PNO method (including electron correlation). The influence of polarization functions and electronic correlation on the shape of the potential curves of the two states is discussed. The calculated equilibrium geometries agree very well with experiment. The results for transition energies (e.g.  $^3B_1 \rightarrow ^1A_1$  excitation energy = 9.2 kcal/mole, total binding energy = 187 kcal/mole) are more reliable than the existent experimental values.

**Key words:** Methylene    Electronic structure    Pair correlation energies    Excitation energy.

#### 1. Introduction

In the past years the interest in the properties and the chemistry of the methylene (or "carbene") radical [1–3] has been strongly increased since this radical is an important short-living intermediate in many organic reactions. It has two low-lying states, a triplet and a singlet, which are energetically very close and behave differently in chemical reactions [4]. Thus, methylene is a good example to study the different behaviour of singlet and triplet radicals.

Most of the properties of the two states of methylene are well established today. From spectroscopical measurements (optical spectra [5–8] and ESR [9, 10]) and from quantum chemical calculations (compare Table 1) it is quite certain that the triplet state is the ground state. The equilibrium angles are rather different ( $102.4^\circ$  for the singlet [6] and  $136 \pm 8^\circ$  for the triplet state [8–10]), whereas the C–H equilibrium distances are almost equal in the two states [5–8]. The value  $\Delta E_{3 \rightarrow 1}$  of the triplet-singlet separation, however, is rather uncertain: From optical spectra no information about  $\Delta E_{3 \rightarrow 1}$  is available because there is no intercombination between the singlet and the triplet spectrum. Herzberg estimates an upper limit for  $\Delta E_{3 \rightarrow 1}$  of 23 kcal/mole [7]. Thermochemical experiments (photolysis of ketene [11, 12]) yield much smaller values like 2.7 kcal/mole [11] and 1–2 kcal/mole [12]. Deactivation of the singlet and triplet species, on the other hand, leads to a lower bound of  $kT$  (for 98°K), i.e. to  $\sim 0.6$  kcal/mole [4].

Table 1. Previous calculations for the triplet-singlet separation energy  $\Delta E_{3 \rightarrow 1} = E(^1A_1) - E(^3B_1)$  of methylene

$E(^3B_1)^a$ [a.u.]	$\Delta E_{3 \rightarrow 1}$ [kcal/mole]	Polarization functions	Method	Ref.
- 38.436	40		SCF	[19]
38.8696	37	-	SCF	[19]
- 38.892	40.6		SCF	[20]
- 38.895	57.7		SCF	[13]
	33.5		SCF	[21]
- 38.9078	34.6		SCF	[22]
38.9123	24.9	<i>d</i> at C-atom	SCF	[14]
- 38.913	32.1		SCF	[23]
38.9202	23.9	<i>d</i> at C-atom	SCF	[24]
38.921	27.0	C-H bond funct.	SCF	[25]
38.904	24.5		limited CI	[26]
38.908	24		limited CI	[27]
38.915	32		VB-CI	[20]
38.9248	20		limited CI	[22]
38.9598	11.5	<i>d</i> at C-atom	GVB CI	[24]
38.9822	22.3		INO CI	[23]
- 39.0121	13.8	<i>d</i> at C-atom	INO CI	[14, 15]

<sup>a</sup> The geometry is not optimized in all calculations. We took always the lowest energy value for the  $^1B_1$  state.

All previous theoretical calculations yield much higher values for  $\Delta E_{3 \rightarrow 1}$  in the region from 40 to 10 kcal/mole, as it can be seen from Table 1. A critical analysis of these calculations shows the following characteristic features: Apart from the very early SCF-calculation of Krauss [13] which was performed only for an angle of  $180^\circ$  for the triplet state all SCF-calculations that do not use polarization functions (i.e. *d*-functions at the C-atom and *p*-functions at the H-atoms) yield values for  $\Delta E_{3 \rightarrow 1}$  between 40 and 32 kcal/mole. The inclusion of polarization functions lowers  $\Delta E_{3 \rightarrow 1}$  appreciably, to about 25 kcal/mole. The influence of the electronic correlation is still more important: CI treatments without polarization functions give values between 25 and 20 kcal/mole, whereas the addition of polarization functions lowers this value to about 12 kcal/mole. But even this value is appreciably higher than the experimental results.

The question is still open, whether the rather low experimental estimates or the much higher theoretical results are more reliable. The latter have been obtained with severe approximations in the calculation of the correlation energy and with quite modest basis sets, except for the most recent results of Bender and coworkers [14, 15]. But there are also arguments indicating that the thermochemical estimates might be too small [14].

In this paper we present the results of IEPA-PNO calculations for the two lowest states of methylene in order to clarify the discrepancy concerning the value of  $\Delta E_{3 \rightarrow 1}$ . In the first step we have performed quite accurate SCF-calculations for both states with an extended basis set with and without polarization functions (Sections 2 and 3). Subsequently, CI calculations in the "independent electron pair approach" (IEPA) are performed to get reliable estimates of the

valence shell correlation energy for both states (Section 4). Previous calculations of correlation energies of small hydrides within the IEPA-scheme both for closed shell and open shell states have shown that we can get excitation energies that are accurate to a few kcal/mole [16-18].

In Section 5 we collect some further calculated properties of methylene.

## 2. Description of the Orbital Basis

For all the calculations reported here we have used an orbital basis set of contracted Gaussian lobe functions.  $p$ -functions are constructed from two lobes with equal exponents  $\eta$ , fixed coefficients (1.0 and -1.0), and a distance  $d$  between them such that  $\frac{d}{2}\sqrt{\eta} = 0.1$ .  $d$ -functions are constructed in the same way, except for  $d_{zz}$  for which the distance between the two positive lobes is somewhat larger ( $\frac{d}{2}\sqrt{\eta} = 0.3$ ) and the negative torus is discarded since it can be simulated by the  $s$ -functions at the same center.

Table 2. Gaussian basis set for CH<sub>2</sub> calculations

No.	Symmetry	Coefficient	$d/2[a_0]$	Exponent
1	$s_C$	0.00088 0.00716 0.03864 0.15492 0.45262 0.85724	0*	9470.52 1397.56 307.539 84.5419 26.9117 9.4090
2		1	0	3.50002
3		1	0	1.2322
4		1	0	0.423084
5		1	0	0.147699
6		1	0	0.048957
7-9	$p_C$	0.1752 1.0928 3.6730	0.020 0.042 0.075	25.3655 5.77636 1.7873
10-12		1	0.123	0.655774
13-15		1	0.199	0.25279
16-18		4.2316 0.1738	0.322 0.606	0.096693 0.027189
19, 20	$s_H$	0.02448 0.18300 0.82288	0	33.6444 5.05796 1.1468
21, 22		1	0	0.321144
23, 24		1	0	0.101309
25, 26	$p_{\sigma, H}$	1	0.079	1.6
27, 28		1	0.158	0.4
29-33	$d_C$	1	0.097	1.06803
34-38		1	0.176	0.321144
39-42	$p_{\pi, H}$	1	0.124	0.65

\*  $d/2$  is the distance of the lobes from the corresponding center (in  $a_0$ ).

For the C-atom we started from the  $11s, 7p$  basis set given by Huzinaga [28] contracted to 6  $s$ -groups and 4  $p$ -groups with (6, 1, 1, 1, 1, 1) and (3, 1, 1, 2) contraction, respectively. For the H-atoms a  $5s$  Huzinaga [28] basis with (3, 1, 1) contraction was used. All exponents and contraction coefficients are given in Table 2. Some further reoptimizations within this set (exponents, contraction pattern, and contraction coefficients) give minor corrections to the SCF-energies (in the order of 0.001 a.u. or less) and can be neglected.

This conventional basis set of approximately triple zeta quality has been enlarged by the addition of polarization functions: Two sets of five  $d$ -functions at the C-atom with exponents of 1.06803 and 0.321144, two  $p_\sigma$  groups at each H-atom with exponents of 1.6 and 0.4, and one  $p_\pi$  group in either direction at the H-atoms with an exponent of 0.65. All these exponents are chosen reasonably (compare [29]), but not fully optimized since the total energy is fairly insensitive to minor changes of them. Furthermore, a full optimization for a particular geometry and state may destroy the balance of the basis set.

### 3. SCF-Calculations

According to elementary molecular orbital theory  $\text{CH}_2$  has the electronic configuration

$$1\sigma_g^2 2\sigma_g^2 1\sigma_u^2 1\pi_u^2$$

in its linear conformation. Because of the twofold degeneracy of the  $1\pi_u$  orbital this configuration leads to three low lying electronic states

$$^3\Sigma_g^-, \quad ^1A_g, \quad ^1\Sigma_g^+.$$

When the HCH bond angle deviates from  $180^\circ$  the degeneracy of the  $1\pi_u$  orbital is removed and the  $^1A_g$  state splits into an  $^1A_1$  and a  $^1B_1$  state. Consequently, we get the following correlation between the states of the linear and the bent conformation of  $\text{CH}_2$

$$1\sigma_g^2 2\sigma_g^2 1\sigma_u^2 1\pi_u^2 \left\{ \begin{array}{ll} ^1\Sigma_g^+ \rightarrow ^1A_1^* & 1a_1^2 2a_1^2 1b_2^2 1b_1^2 \\ & ^1B_1 \quad 1a_1^2 2a_1^2 1b_2^2 3a_1 1b_1 \\ ^1A_g \rightarrow \begin{cases} ^1B_1 \\ ^1A_1 \end{cases} & \begin{cases} 1a_1^2 2a_1^2 1b_2^2 3a_1 1b_1 \\ 1a_1^2 2a_1^2 1b_2^2 3a_1^2 \end{cases} \\ ^3\Sigma_g^- \rightarrow ^3B_1 & 1a_1^2 2a_1^2 1b_2^2 3a_1 1b_1 \end{array} \right.$$

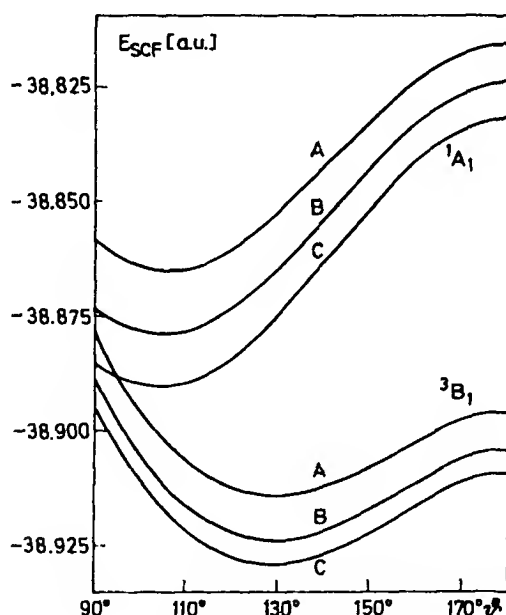
According to Walsh's rules [30] the lowest  $^1A_1$  state should be bent, the  $^3B_1$  state, however, is expected to be linear ( $=^3\Sigma_g^-$ ) or almost linear. Due to the quasi-degeneracy of the  $3a_1$  and  $1b_1$  orbitals there should be a strong configuration mixing of the two  $^1A_1$  states for large bond angles. Consequently, the SCF-approximation cannot be very good for the  $^1A_1$  state at large angles.

As first step in our IEPA-treatment of the  $\text{CH}_2$  radical we have performed restricted SCF-calculations for the two lowest states ( $^3B_1$  and  $^1A_1$ ) for various angles and fixed C-H distance  $R = 2.10 a_0$ . A Roothaan-McWeeny type SCF-programme has been used for the open- and closed-shell SCF-calculations [31-33]. To investigate the influence of the polarization functions on the two states and on the value of  $\Delta E_{3-1}$  we have used three different basis sets:

Basis A: without polarization functions; this basis consists of the first 24 groups of Table 2.

Table 3. SCF-energies for  $\text{CH}_2$  with different basis sets.  $R = 2.1 a_0$  (all values negative, in a.u.)

Angle $\theta$	State	Basis A	Basis B	Basis C
90°	$^1A_1$	38.85868	38.87386	38.88582
	$^3B_1$	38.87947	38.88956	38.89563
105°	$^1A_1$	38.86549	38.87893	38.89060
	$^3B_1$	38.90215	38.91187	38.91755
120°	$^1A_1$	38.86057	38.87326	38.88408
	$^3B_1$	38.91260	38.92237	38.92772
135°	$^1A_1$	38.84776	38.85970	38.86967
	$^3B_1$	38.91355	38.92323	38.92835
150°	$^1A_1$	38.83241	38.84298	38.85206
	$^3B_1$	38.90794	38.91715	38.92221
165°	$^1A_1$	38.82042	38.82963	38.83738
	$^3B_1$	38.89994	38.90846	38.91351
180°	$^1A_1 = ^1\Delta_g$	38.81602	38.82447	38.83162
	$^3B_1 = ^3\Sigma_g$	38.89583	38.90402	38.90900
$\Delta E_{3 \rightarrow 1}$ (a.u.)		0.0487	0.0450	0.0385
$\Delta E_{3 \rightarrow 1}$ (kcal/mole)		30.6	28.2	24.2
$\theta_0$	$^1A_1$	106.2°	104.6°	103.9°
$\theta_0$	$^3B_1$	129.1°	128.9°	128.5°

Fig. 1. SCF-energies of the two lowest states of  $\text{CH}_2$  calculated with different basis sets.  $R_{\text{CH}} = 2.1 a_0$ 

Basis B: with two  $p_\sigma$ -functions at each H-atom; this basis consists of the first 28 groups of Table 2.

Basis C: two sets of  $d$ -functions at C, two  $p_\sigma$  and one  $p_\pi$  at each H-atom; full basis of Table 2.

The results of these SCF-calculations are collected in Table 3 and Fig. 1 and can be analyzed in the following way:

1) The qualitative features of the previous results in Table I are confirmed. Our calculation  $A$  coincides roughly with previous calculations without polarization functions [21–23] as far as the ground state energy and  $\Delta E_{3 \rightarrow 1}$  are concerned. The inclusion of the  $p_\sigma$ -functions (basis B) lowers  $\Delta E_{3 \rightarrow 1}$  by 2.4 kcal/mole, resulting in energy values which are rather similar to those of Meyer [25] who had used "bond functions" located in the C-H bonds. Finally, a further reduction of  $\Delta E_{3 \rightarrow 1}$  by 4.0 kcal/mole is achieved by the inclusion of  $d_C$  and  $p_{\pi,H}$  functions (basis C) giving  $\Delta E_{3 \rightarrow 1} = 24.2$  kcal/mole. This value is almost identical to those of Bender *et al.* [14] and Hay *et al.* [24] though these authors use much poorer  $s$ - and  $p$ -basis sets and get less good ground state energies.

2) This overall behaviour can be understood by looking at the details in Table 3. The polarization functions, particularly  $d_C$  and for small angles also  $p_{\pi,H}$ , lower the  $^1A_1$  state much more than the  $^3B_1$  state. The reason is obvious: The inclusion of the polarization functions makes the basis more flexible as far as the  $3a_1$  (lone pair) orbital is concerned, but does not affect the  $1b_1$  orbital which is antisymmetrical with respect to the molecular plane. Since the  $3a_1$  orbital is doubly occupied in the  $^1A_1$  state and singly occupied in the  $^3B_1$  state, the gain in energy for  $^1A_1$  is approximately twice as large as for  $^3B_1$ .

3) For large angles the  $3a_1$  orbital is similar to a  $\pi_z$  orbital which is little affected by polarization functions (just by  $p_{\pi,H}$ ). For small angles, however, it is a  $sp^2$  hybrid pointing away from the carbon atom and is influenced very strongly by the  $d_C$  functions. Therefore, the inclusion of polarization functions favours smaller bond angles; this effect again is more remarkable for the  $^1A_1$  state ( $-2.3^\circ$ ) than for the  $^3B_1$  state ( $-0.6^\circ$ ).

We conclude that the behaviour of the SCF-potential curves and the value of  $\Delta E_{3 \rightarrow 1}$  depends strongly on the polarization functions. The inclusion of them is even more important than the use of an extended  $s$ - and  $p$ -basis though this might give lower SCF-energies.

#### 4. Pair Correlation Energies in $\text{CH}_2$

Our SCF-value for  $\Delta E_{3 \rightarrow 1}$  of 24.2 kcal/mole is still much higher than the experimental estimates [4, 7, 11, 12]. This is due to the neglect of the electronic correlation which lowers the total energy of the  $^1A_1$  state with respect to the  $^3B_1$  state such that  $\Delta E_{3 \rightarrow 1}$  becomes smaller. Such a behaviour can be expected since the number of doubly occupied SCF-orbitals is higher in the  $^1A_1$  state than in the  $^3B_1$  state. Furthermore, for both states the correlation energy changes appreciably with the bond angle and affects noticeably the shape of the potential curves.

In our calculations we take care of the electronic correlation by means of the IEPA-PNO method ("independent electron pair approach with pair natural orbitals"). Since this method has been reviewed recently for closed and open shell states [34–36] we do not go into the details here. In this method the total correlation energy of a molecule is approximated by a sum over pair correlation energies  $\epsilon_{ij}$ :

$$E_{\text{corr}} \approx E_{\text{corr}}(\text{IEPA}) = \sum_i \epsilon_{ii} + \sum_{i < j} ({}^1\epsilon_{ij} + {}^3\epsilon_{ij})$$



where  $i$  and  $j$  denote the occupied SCF-orbitals. One distinguishes between *intrapair* correlation energies  $\epsilon_{ii}$  and singlet and triplet *interpair* correlation energies  ${}^1\epsilon_{ij}$  and  ${}^3\epsilon_{ij}$  (for  $i \neq j$ , singlet or triplet coupling of the orbitals  $i$  and  $j$ , respectively).

The  $\epsilon_{ij}$  are calculated independently from each other by means of a natural expansion [37] in terms of "pair natural orbitals" [38] (PNO's). With basis sets similar to those used in our calculation one generally gets 80–90% of the individual pair correlation energies.

Unfortunately,  $E_{\text{corr}}$  (IEPA) is only an approximation to the total correlation energy  $E_{\text{corr}}$ , the deviation  $E_{\text{corr}} - E_{\text{corr}}(\text{IEPA})$  may have either sign. The error is mainly due to the fact that the  $\epsilon_{ij}$  are calculated independently from each other where certain coupling terms are neglected ("additivity error"). Previous calculations, however, have shown that the IEPA results for properties like binding energies, excitation energies, force constants etc. generally are much more reliable than SCF-results, at least for small atoms [17] and first row hydrides [39].

In order to minimize these additivity errors, to save computation time, and to simplify the interpretation of the results we use localized SCF-orbitals

according to the localization criterion of Boys [40] — and calculate pair correlation energies of localized pairs. In the closed shell  ${}^1A_1$  state of  $\text{CH}_2$  all four doubly occupied SCF-orbitals can be localized and the SCF-determinant can be written

$$\Psi({}^1A_1) = |1a_1^2 2a_1^2 1b_2^2 3a_1^2| = |K^2 l^2 r^2 n^2|$$

where  $K, l, r, n$  denote the carbon  $K$ -shell, the left and right C-H bond orbitals, and the lone pair orbital, respectively. In the  ${}^3B_1$  state, however, a complete localization is not possible since one cannot mix doubly and singly occupied SCF-orbitals. So we are left with

$$\Psi({}^3B_1) = |1a_1^2 2a_1^2 1b_2^2 3a_1 1b_1| = |K^2 l^2 r^2 3a_1 1b_1|$$

The singly occupied  $3a_1$ -orbital is almost identical with the lone pair orbital  $n$ .

Tables 4 and 5 contain our results for the valence shell pair correlation energies calculated with basis C as functions of the bond angle  $\vartheta$  for fixed C-H distance  $R = 2.10 a_0$ .  $E_C$  is the sum of the  $\epsilon_{ij}$ , i.e. the IEPA estimate of the valence shell correlation energy. Since the  $K$ -shell and  $K$ - $L$ -intershell contributions are almost independent of  $R$  and  $\vartheta$  [39] we neglect them and expect  $E = E_{\text{SCF}} + E_C$  to represent the variation of the total energy with  $\vartheta$  very well.

The interpretation of the results in the Tables 4 and 5 is rather simple and illustrative.

1) The intrapair correlation energies  $\epsilon(l^2)^1$  for the C-H bonds are almost independent of  $\vartheta$ .  $\epsilon(n^2)$ , however, is rather sensitive to changes in  $\vartheta$  because of the change in hybridization and of the quasi-degeneracy of the  $3a_1$  and  $1b_1$  orbitals for large angles.

2) The interpair correlation energies  $\epsilon_{ij}(i \neq j)$  depend very strongly on the differential overlap of the localized orbitals  $i$  and  $j$ . Thus,  $\epsilon(l, r)$  becomes smaller — in absolute value — with increasing  $\vartheta$ . Similarly,  $|\epsilon(l, n)|$  for  ${}^1A_1$  and

<sup>1</sup> To avoid too much lower indices we adopt the notation  $\epsilon(i^2)$ ,  $\epsilon(i, j)$  instead of  $\epsilon_{ii}$ ,  $\epsilon_{ij}$  if the orbitals  $i$  and  $j$  are explicitly specified. Furthermore:  $\epsilon(i, j) = {}^1\epsilon(i, j) + {}^3\epsilon(i, j)$ .

Table 4 Angular dependence of the valence shell pair correlation energies for the  $^3B_1$  state of  $\text{CH}_2$  (all energy values in a.u.,  $R = 2.1 a_0$ )

		90°	105°	120°	135°	150°	165°	177°
$L_{SC1}$	$n^a$	38.89563	-38.91755	-38.92772	-38.92835	-38.92221	-38.91351	-38.90922
$\epsilon(l^2)$	2	0.02992	0.02956	0.02927	-0.02909	-0.02897	-0.02883	-0.02871
$\epsilon(l, r)^b$	1	0.01453	0.01304	-0.01206	-0.01138	-0.01087	-0.01052	-0.01037
$\epsilon(l, 3a_1)^b$	2	0.01330	0.01363	0.01430	-0.01532	-0.01691	-0.01907	-0.02044
$\epsilon(l, 1b_1)^b$	2	0.01758	0.01779	0.01813	0.01861	0.01927	-0.02015	-0.02068
${}^1\epsilon(3a_1, 1b_1)^c$	1	0.00776	0.00798	0.00820	0.00844	0.00870	0.00892	0.00925
$E_c$		0.14389	0.14298	0.14366	-0.14586	0.14987	0.15554	-0.15928
$E_{SC1} + E_c$		39.03952	39.06053	39.07138	39.07421	39.07208	39.06905	-39.06850

<sup>a</sup> Number of equivalent pairs with the same  $r_{ij}$ .<sup>b</sup> Sum of  ${}^1\epsilon_{ij}$  and  ${}^1\epsilon_{ji}$ .<sup>c</sup> In this case only the  ${}^1\epsilon_{ij}$  occurs [36].Table 5 Angular dependence of the valence shell pair correlation energies for the  ${}^1A_1$  state of  $\text{CH}_2$  (all energy values in a.u.,  $R = 2.1 a_0$ )

		90°	105°	120°	135°	150°
$L_{SC1}$	$n^a$	38.88582	38.89060	-38.88408	-38.86967	38.85206
$\epsilon(l^2)$	2	0.03258	0.03270	-0.03277	-0.03276	-0.03251
$\epsilon(l, r)^b$	1	-0.01915	0.01792	-0.01727	-0.01704	-0.01711
$\epsilon(n^2)$	1	0.03920	-0.03906	-0.03825	-0.03688	-0.03541
$\epsilon(l, n)^b$	2	0.02378	-0.02377	-0.02409	-0.02471	-0.02563
$E_c$		-0.17107	0.16992	0.16924	-0.16886	-0.16880
$E_{SC1} + E_c$		-39.05689	-39.06052	-39.05332	-39.03853	-39.02086

<sup>a</sup> Number of equivalent pairs with the same  $r_{ij}$ .<sup>b</sup> Sum of  ${}^1\epsilon_{ij}$  and  ${}^1\epsilon_{ji}$ .

$|e(l, 3a_1)|$  and  $|{}^3e(3a_1, 1b_1)|$  for  ${}^3B_1$  become smaller for decreasing angle since the  $3a_1$  orbital (the lone pair in  ${}^1A_1$ ) is pushed away from the C-atom changing from a  $p_z$  orbital at  $180^\circ$  to a  $sp^2$  hybrid at  $120^\circ$ .  $\epsilon(l, 1b_1)$ , however, depends less significantly on  $\vartheta$  since the distance of the centers of gravity of the  $l$  and  $1b_1$  orbitals does not depend on  $\vartheta$ .

3) Since the  $1b_1$  orbital is not occupied in the SCF-determinant for the  ${}^1A_1$  state it is still available for excitation in the CI-calculation. Consequently, the absolute values of  $\epsilon(l^2)$  and  $\epsilon(l, r)$  are appreciably larger in the  ${}^1A_1$  state than in the  ${}^3B_1$  state.

From these details the overall behaviour can be understood quite easily.  $|E_c|$  is much larger in the  ${}^1A_1$  state than in the  ${}^3B_1$  state. This is due to the large value of  $|e(n^2)|$  which has no counterpart in the  ${}^3B_1$  state and to the unoccupied  $1b_1$  orbital causing larger values of  $|e(l^2)|$  and  $|e(l, r)|$  for the  ${}^1A_1$  state. At the energy minima near  $105^\circ$  and  $135^\circ$ , respectively, the difference in the correlation energies is about  $0.024 \text{ a.u.} = 15 \text{ kcal/mole}$ . Thus, the influence of correlation lowers  $\Delta E_{3-1}$  from 24 to about 9 kcal/mole.

The angular dependence of  $E_c$  — see Fig. 2 — can be understood as well in terms of the  $\epsilon_{ij}$ . In the  ${}^1A_1$  state the changes in  $\epsilon(l, r)$  and  $\epsilon(l, n)$  cancel such that

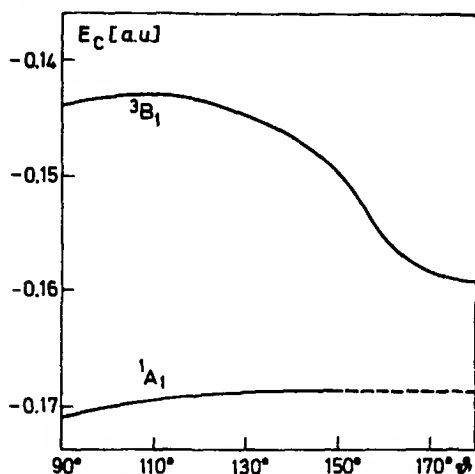


Fig. 2. Angular dependence of the valence shell correlation energy  $E_C$  of the two lowest states of  $\text{CH}_2$ .  $R_{\text{CH}} = 2.1 a_0$

$E_C$  is almost independent of  $\theta$ . For angles smaller than  $105^\circ$  the increase in  $|\epsilon(l, r)|$  dominates causing an increase in  $|E_C|$  and a decrease in the equilibrium bond angle. For the  $^3B_1$  ground state, however, a strong increase of  $|E_C|$  with increasing  $\theta$  results such that the well depth — i.e. the energy difference  $E(180^\circ) - E(135^\circ)$  — is lowered appreciably from 12 kcal/mole in the SCF-approximation to 3.5 kcal/mole with correlation. The equilibrium bond angle is enlarged from  $128.5^\circ$  to  $134.2^\circ$ .

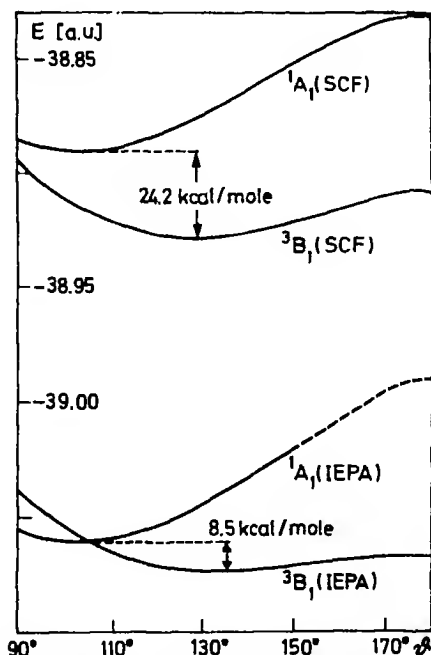


Fig. 3. SCF- and IEPA-potential energy curves for the two lowest states of  $\text{CH}_2$ .  $R_{\text{CH}} = 2.1 a_0$

In Fig. 3 the potential curves  $E(\vartheta)$  are given for fixed  $R$  both in SCF-approximation and with correlation.

### 5. Molecular Properties of $\text{CH}_2$

To get a last correction to the value of  $\Delta E_{3 \rightarrow 1}$  we have to take into account that the C-H distance for both states can be different. Therefore, we have calculated the SCF- and valence shell correlation energies for both states for a symmetrical variation of  $R$  at fixed bond angles (Table 6).

Table 6 Dependence of SCF- and valence shell correlation energies in the  $^3B_1$  and  $^1A_1$  states of  $\text{CH}_2$  on the C-H distance ( $R = R_1 = R_2$ )

		$R = 1.9 a_0$	$R = 2.0 a_0$	$R = 2.1 a_0$	$R = 2.2 a_0$
$^1A_1$ 102.5	$E_{\text{SCF}}$	38.87889	38.88907	- 38.89059	- 38.88585
	$E_c$		0.16773	- 0.17008	- 0.17216
	$E_{\text{SCF}} + E_c$		39.05680	39.06067	- 39.05801
$^3B_1$ 135	$E_{\text{SCF}}$	38.92441	- 38.93063	- 38.92835	- 38.91979
	$E_c$	0.14236	0.14406	0.14586	- 0.14763
	$E_{\text{SCF}} + E_c$	39.06677	- 39.07469	39.07421	39.06742

Table 7 Molecular properties of  $\text{CH}_2$

Property	SCF	II:PA <sup>a</sup>	Exp.
$^3B_1$ ground state			
total energy (a.u.)	- 38.93078	- 39.07539	
$R_0$ (Å)	1.069	1.081	(1.029 [5, 7]) 1.078 [8]
$\vartheta_0$	128.5	134.2	136 ± 8 [9, 10]
well depth (kcal/mole)	12.2	3.5	
$F_R$ (mdyn/Å)	6.28	5.85	
$F_\vartheta$ (mdyn/Å)	0.45	0.28	
binding energy (kcal/mole)	155	187	180 ± 5 [42] <sup>b</sup> (208 [43]) <sup>b</sup>
$^1A_1$ excited state			
total energy (a.u.)	38.89090	- 39.06070	
$R_0$ (Å)	1.095	1.116	1.11 <sub>1</sub> [6]
$\vartheta_0$	103.9	102.5	102.4 [6]
well depth (kcal/mole)	37.1	(43)	
$F_R$ (mdyn/Å)	5.44	5.08	
$F_\vartheta$ (mdyn/Å)	0.60	0.55	
$^3B_1 \rightarrow ^1A_1$ crossing angle	80	105	
$1E_{3 \rightarrow 1}$ (kcal/mole)	25.1	9.2	0. . . 23 <sup>c</sup>
$^3B_1 \rightarrow ^1A_1$ (135°) (kcal/mole)	37	22	
$^3B_1 \rightarrow ^1A_1$ (180°) (kcal/mole)	49	(50)	
$^1A_1$ (102.5°) $\rightarrow$ $^1\Sigma_u^+$ (kcal/mole)		(69)	79-86 [6]

Values in parentheses are less certain.

<sup>a</sup> Only valence shell correlation included.

<sup>b</sup> Experimental heat of formation.

<sup>c</sup> Compare discussion in the introduction.

One observes that  $|E_C|$  increases almost linearly with increasing  $R$ . This leads to larger C-H equilibrium distances as compared to the SCF-values, to larger anharmonicities of the potential curves, and to lower harmonic force constants (compare Table 7). The value for  $\Delta E_{3 \rightarrow 1}$  is changed slightly to 9.2 kcal/mole. The  $R$ -dependence of  $|E_C|$  is mainly due to the increase of  $|\epsilon(l^2)|$  with  $R$  and has been discussed in earlier papers of this series [39].

Table 7 contains values for some further molecular properties calculated for both states (with basis set C) and compared with experimental results. The calculated equilibrium distances and angles are in good agreement with experiments, the IEPA values generally are slightly better than the SCF-values. The only discrepancy occurs for  $R_0$  in the  $^3B_1$  state. Herzberg's first value of 1.029 Å [5, 7] disagrees completely with our and other theoretical results [41]. This seems to be a consequence of the first incorrect interpretation of the triplet spectrum of  $\text{CH}_2$  [5]. Indeed, Herzberg and Johns later have shown [8] that the triplet spectrum is compatible with a bent  $^3B_1$  ground state and a C-H distance of 1.078 Å. This value is supported by our calculation as well as by McLaughlin *et al.* [41], particularly if one takes into account that the IEPA generally leads to slightly too large equilibrium distances.

The harmonic force constants  $F_R$  and  $F_\beta$  for the symmetrical stretching and the bending vibration are calculated with neglect of the small coupling term. The results agree with those of McLaughlin *et al.* [41] for the  $^3B_1$  state with the exception that these authors do not find the decrease in  $F_R$  after the inclusion of correlation. Since the angular dependence of the energies is very anharmonic (see Fig. 3), particularly for the  $^3B_1$  state, one should be careful not to calculate the vibrational levels just from  $F_\beta$ .

The second root in the IEPA-CI for the lone pair of the  $^1A_1$  state at 180° gives an estimate for the energy of the excited  $^1A_1^+$  state which is expected to be linear ( $=^1\Sigma_g^+$ ). The energy difference for the excitation  $^1A_1(102.5^\circ) \rightarrow ^1\Sigma_g^+$  is estimated to 69 kcal/mole which agrees with the band observed by Herzberg and Johns [6] between 3330 and 3625 Å (86–79 kcal/mole) and tentatively assigned to a  $^1A_1 \rightarrow ^1\Sigma_g^+$  transition.

The total binding energy for the  $^3B_1$  ground state of  $\text{CH}_2$  is changed appreciably by correlation from 155 kcal/mole in the SCF-approximation to 187 kcal/mole with IEPA. This value has to be reduced by 10 kcal/mole — the estimate for the zero point vibration energy — to be comparable with experimental heats of formation as given by Herzberg [7] (178 kcal/mole) and Prophet [42] ( $180 \pm 5$  kcal/mole). The value recommended by Gaydon [43] (208 kcal/mole) is far off and seems to be completely wrong.

We have to add some remarks concerning the accuracy of our calculations and of the predicted properties. The main error sources in the SCF- and IEPA-calculations are: limitation of the Gaussian basis set, additivity error and neglect of higher substituted configurations in IEPA, neglect of  $K$ -shell correlation. These errors can have different signs and cancel partly, especially if one is interested in transition properties. A detailed analysis of the possible errors occurring at the equilibrium geometries of both states leads us to the following estimates for  $\Delta E_{3 \rightarrow 1}$

SCF-limit	$22 \pm 2$ kcal/mole
final value	$9 \pm 3$ kcal/mole

which agree fairly well with the estimates given by Bender *et al.* [14]. IEPA calculations with comparable basis sets on the carbon atom [17] yield transition energies (excitation energies, ionization energy, and electron affinity) with errors of the same order of 1–3 kcal/mole and support our above error estimate.

The situation is less favourable for the  $^1A_1$  state at large angles where the mixing of the configurations  $\dots 3a_1^2$  and  $\dots 1b_1^2$  is large and a one-determinant SCF-wavefunction is a bad approximation. Since our IEPA-method starts from an one-determinant SCF-approximation we expect larger errors in the potential curve of the  $^1A_1$  state for large  $\theta$ . The corresponding properties in Table 7 are given in parentheses, the potential curves in the Figs. 2 and 3 are given with dashed lines.

*Acknowledgment.* The author is indebted to Prof. Dr. W. Kutzelnigg and Dr. R. Ahlrichs (Karlsruhe) and to Dr. M. Jungen (Basel) for their interest in this work and numerous discussions on this subject. The calculations have been performed at the „Rechenzentrum der Universität Karlsruhe“

## References

1. Gaspar, P. P., Hammond, G. S.: In: Kirmse, W. (Ed.): Carbene chemistry. New York: Academic Press 1964
2. Closs, G.: Topics Stereochem. **3**, 198 (1968)
3. Bethell, D.: Advan. Phys. Org. Chem. **7**, 153 (1969)
4. Braun, W., Bass, A. M., Pilling, M.: J. Chem. Phys. **52**, 5131 (1970)
5. Herzberg, G.: Proc. Roy. Soc. (London) A **262**, 291 (1961)
6. Herzberg, G., Johns, J. W. C.: Proc. Roy. Soc. (London) A **295**, 107 (1966)
7. Herzberg, G.: Electronic spectra and electronic structure of polyatomic molecules. Princeton: van Nostrand 1967
8. Herzberg, G., Johns, J. W. C.: J. Chem. Phys. **54**, 2276 (1971)
9. Wasserman, F., Kuck, V. J., Hutton, R. S., Yager, W. A.: J. Am. Chem. Soc. **92**, 7491 (1970)
10. Wasserman, F., Yager, W. A., Kuck, V. J.: Chem. Phys. Letters **7**, 409 (1970)
11. Halberstadt, M. L., McNesby, J. R.: J. Am. Chem. Soc. **89**, 3417 (1967)
12. Carr, R. W., Fider, T. W., Topor, M. G.: J. Chem. Phys. **53**, 4716 (1970)
13. Krauss, M.: J. Res. Natl. Bur. Stand. **68 A**, 635 (1960)
14. Bender, C. F., Schaefer III, H. F., Franceschetti, D. R., Allen, L. C.: J. Am. Chem. Soc. **94**, 6888 (1972)
15. Schaefer III, H. F.: The electronic structure of atoms and molecules, p. 309. Reading: Addison-Wesley 1972
16. Staemmler, V., Jungen, M.: Chem. Phys. Letters **16**, 187 (1972)
17. Staemmler, V.: Unpublished results on atoms.
18. Driessler, F., Ahlrichs, R., Staemmler, V., Kutzelnigg, W.: To be published
19. Lathan, W. A., Hehre, W. J., Curtiss, L. A., Pople, J. A.: J. Am. Chem. Soc. **93**, 6377 (1971)
20. Harrison, J., Allen, L. C.: J. Am. Chem. Soc. **91**, 807 (1969)
21. Del Bene, J. E.: Chem. Phys. Letters **9**, 68 (1971)
22. Chu, S. Y., Siu, K. Q., Hayes, F. F.: J. Am. Chem. Soc. **94**, 2969 (1972)
23. O'Neill, S. V., Schaefer, III, H. F., Bender, C. F.: J. Chem. Phys. **55**, 162 (1971)
24. Hay, P. J., Hunt, W. J., Goddard III, W. A.: Chem. Phys. Letters **13**, 30 (1972)
25. Meyer, W.: Dissertation, München 1968, unpublished
26. Foster, J. M., Boys, S. F.: Rev. Mod. Phys. **32**, 305 (1960)
27. Harrison, J. F.: J. Am. Chem. Soc. **93**, 4112 (1971)
28. Huzinaga, S.: J. Chem. Phys. **42**, 1293 (1965)
29. Rothenberg, S., Schaefer III, H. F.: J. Chem. Phys. **54**, 2764 (1971)
30. Walsh, A. D.: J. Chem. Soc. **1953**, 2260
31. Roothaan, C. C. J.: Rev. Mod. Phys. **23**, 69 (1951)
32. Roothaan, C. C. J.: Rev. Mod. Phys. **32**, 179 (1960)

33. McWeeny, R.: *Molecular orbitals in chemistry, physics, and biology*, p. 305. New York: Academic Press 1964
34. Kutzelnigg, W.: *Molecular calculations including electron correlation*. In: Clementi, E. (Ed.): *Selected topics in molecular physics*, p. 91. Weinheim: Verlag Chemie 1972
35. Kutzelnigg, W.: *Electron correlation and the electron pair theories*, NATO Summer School, Ramsau 1971
36. Staemmler, V., Jungen, M.: To be published
37. Löwdin, P. O., Shull, H.: *Phys. Rev.* **101**, 1730 (1956)
38. Edmiston, C., Krauss, M.: *J. Chem. Phys.* **45**, 1833 (1966)
- 39a. Ahlrichs, R., Kutzelnigg, W.: *Theoret. Chim. Acta (Berl.)* **10**, 377 (1968)
- 39b. Ahlrichs, R., Kutzelnigg, W.: *Chem. Phys. Letters* **1**, 651 (1968)
- 39c. Jungen, M., Ahlrichs, R.: *Theoret. Chim. Acta (Berl.)* **17**, 339 (1970)
- 39d. Ahlrichs, R.: *Theoret. Chim. Acta (Berl.)* **17**, 348 (1970)
- 39e. Jungen, M.: *Chem. Phys. Letters* **5**, 241 (1970)
- 39f. Géfus, M., Ahlrichs, R., Staemmler, V., Kutzelnigg, W.: *Theoret. Chim. Acta (Berl.)* **21**, 63 (1971)
- 39g. Staemmler, V., Jungen, M.: *Chem. Phys. Letters* **16**, 187 (1972)
- 39h. Géfus, M., Kutzelnigg, W.: *Theoret. Chim. Acta (Berl.)* **28**, 103 (1973)
- 39i. Lischka, H.: To be published
40. Foster, J. M., Boys, S. F.: *Rev. Mod. Phys.* **32**, 300 (1960)
41. McLaughlin, D. R., Bender, C. F., Schaefer III, H. F.: *Theoret. Chim. Acta (Berl.)* **25**, 352 (1972)
42. Prophet, H.: *J. Chem. Phys.* **38**, 2345 (1963)
43. Gaydon, A. G.: *Dissociation energies and spectra of diatomic molecules*. London: Chapman and Hall 1968

Dr. Volker Staemmler  
 Institut für Physikalische Chemie  
 und Elektrochemie der Universität Karlsruhe  
 D-7500 Karlsruhe, Kaiserstraße 12  
 Federal Republic of Germany





# A New Definition of Atomic Charges in Molecules\*

Karl Jug

Institut für Theoretische Chemie der Universität Stuttgart and Department of Chemistry,  
Saint Louis University Saint Louis, Missouri 63156, USA\*\*

Received April 9, 1973

A new definition of atomic charges in molecules is presented which conserves charge and dipole moment. It contains the Mulliken and Löwdin definitions as special cases of zero and first order truncations of commutator expansions. The definition allows for a systematic improvement of charges paralleling the improvement of the basis set in the LCAO approximation. We have tested the definition in thirteen selected diatomics and polyatomics in optimal minimal Slater basis set SCF calculations by means of 4G-level Gaussian expansions. The results suggest that the proposed definition is better than either Mulliken's or Löwdin's definition.

*Key words:* Atomic charges Dipole moments.

## 1. Introduction

In recent years, many attempts have been made to improve the standard population analysis suggested by Mulliken [1] and also by Daudel [2]. They fall essentially in three categories:

i) The partitioning of two-center atomic orbital overlap integrals in atomic parts. The Mulliken definition falls in this category. More sophisticated definitions are those by Löwdin [3], which conserves the dipole moment of a two-center charge distribution and by Christoffersen [4], which uses coefficients of an SCF calculation for weighting.

ii) The use of symmetrically orthogonalized atomic orbitals [5]. This method is essentially underlying the CNDO and related methods by Pople [6].

iii) The definition of regions in space which are attributed exclusively to atoms. This method has been pursued qualitatively by means of force fields by Bader [7] and quantitatively through considerations of noninteracting atoms in molecules by Politzer [8].

The advantages of these methods are largely set off by their disadvantages. The third category is most attractive to the experimental chemist and it also offers theoretically a broader range of application than either the first or the second. In particular, the latter ones cannot be directly applied to single-center orbital sets. In a case of single-center MO's with one non-linear variation parameter, Hartmann and Jug [9] segmented the space in atomic regions for five-membered heterocyclic systems to obtain atomic charges. The problem was here as

\* Presented in part at the E.U. Condon Symposium, Sanibel Island, Florida, 21-27 January 1973.

\*\* Permanent address.

well as in the other methods of this category [7, 8] the definition of the atomic regions and the integration. In general, this problem is of crucial difficulty and the latter methods have been applied so far to linear systems only. The advantage of the second method is that no overlap distributions appear explicitly. Implicitly it needs atomic orbitals from other centers to obtain orthogonalized AO's. No justification has been given so far with regard to the relevance of these admixtures. The first category seems to offer the least difficulty if an atomic basis set is used and consequences of partitioning are investigated properly. This means that a relation to a measurable physical quantity has to be established. Only the Löwdin definition based on the conservation of the dipole moment is of this type.

In previous papers [10, 11] we have laid the ground work for a general definition of the first category and applied the concept to several diatomics. The application was limited to four-orbital expansions of integrals related through commutator equations. This paper extends the application to optimal minimal basis set expansions for thirteen selected diatomics and polyatomics. It also gives proof of the conservation of charge and dipole moment in a general fashion. Tables with atomic net charges,  $\sigma-\pi$  separation, atomic occupation numbers, dipole moments and ionic character of bonds are illustrating the differences between Mulliken, Löwdin and this work's definition of atomic charges.

## 2. The Method

The commutator equation

$$u = [t, x] \quad (2.1)$$

with hermitian and antihermitian operators  $t$ ,  $x$  and  $u$  is equivalent to an infinite expansion of integrals over the above operators in a complete basis set  $\chi$

$$u_{\mu\nu} = \sum_{\chi, \chi'} t_{\mu\chi}(S^{-1})_{\chi\chi'} x_{\chi'\nu} - x_{\mu\chi}(S^{-1})_{\chi\chi'} t_{\chi'\nu} \quad (2.2)$$

with

$$u_{\mu\nu} = \langle \mu | u | \nu \rangle \quad \text{etc.}$$

In truncated expansions the above relationship between the integrals in (2.2) is in general only approximately valid.

In the following, we discuss two cases for which (2.2) is representing an equality in a *finite* expansion.

1. Conservation of charge:  $u=0$ ,  $x=1$ .

If we take as expansion functions  $\chi, \chi'$  only those which are used to define the inverse matrix  $S^{-1}$ , it holds that

$$\begin{aligned} & \sum_{\chi, \chi'}^N t_{\mu\chi}(S^{-1})_{\chi\chi'} S_{\chi'\nu} - S_{\mu\chi}(S^{-1})_{\chi\chi'} t_{\chi'\nu} \\ &= \sum_{\chi}^N t_{\mu\chi} \delta_{\chi\nu} - \sum_{\chi'} \delta_{\mu\chi} t_{\chi'\nu} \\ &= t_{\mu\nu} - t_{\mu\nu} \\ &= 0. \end{aligned} \quad (2.3)$$

2. Conservation of dipole moment:  $u = 0, x = r, t = r$ 

$$\sum_{x, x'}^N r_{\mu x} (S^{-1})_{xx'} r_{x'v} - r_{\mu x} (S^{-1})_{xx'} r_{x'v} = 0. \quad (2.4)$$

The first case demonstrates that charge is conserved no matter what operator is chosen for  $t$ . For the conservation of the dipole moment  $t = r$  is necessary. Combining these two observations we choose  $t = r$  for the charge distribution in order to have it on the same footing as the dipole moment partitioning. This means that the total dipole moment of a distribution can be represented as the sum of a charge transfer moment and hybrid atomic moments. We now write (2.2) in such a way that the purpose of partitioning is more apparent. Consider two orbitals  $\mu$  and  $v$  on two different atomic centers A and B. We rewrite (2.2) for  $u = 0$  and  $t = r$  by using only orbitals on centers A and B in the expansion

$$x_{\mu v} = \frac{Q_{\mu v}}{Q_{\mu v}^2} \left[ \sum_{\mu'}^A F_{\mu'v} x_{\mu\mu'} - \sum_{v'}^B F_{v'\mu} x_{v'v} + \sum_{v' \neq v}^B F_{v'v} x_{\mu v'} - \sum_{\mu' \neq \mu}^A F_{\mu'\mu} x_{\mu'v} \right] \quad (2.5)$$

with

$$F_{\mu v} = \sum_x (S^{-1})_{\mu x} r_{xv}$$

$$Q_{\mu v} = F_{\mu\mu} - F_{vv}.$$

Equation (2.5) presents a two-center integral  $x_{\mu v}$  in terms of single-center distributions  $x_{\mu\mu'}$  and  $x_{v'v}$  on atoms A and B plus additional two-center terms  $x_{\mu v'}$  and  $x_{\mu'\mu}$ . If we expand these latter terms again in terms of single-center distributions we can iteratively obtain a representation of any two-center term of  $x$  by single-center terms of atoms A and B. Since we have already proved that for  $x = 1$  and  $x = r$  we obtain equalities (2.3) and (2.4), it means that in (2.5) we distribute two-center overlap integrals among atoms A and B under conservation of charge and dipole moment. In any polyatomic molecule we define a charge transfer in the direction of an internuclear axis. This corresponds to the dipole moment conservation in this direction. In such a local coordinate system there are no two-center integrals perpendicular to the internuclear axis which would need expansion. If partitioning is defined in a local coordinate system and the unitary transformation to molecular coordinate systems performed properly [6], the rotational invariance is guaranteed. The translational invariance is a simple consequence of (2.5) and (2.3). The two-center terms of any other physical quantity  $x$ , which commutes with the dipole moment, can be partitioned iteratively in single-center distributions according to (2.5). Such other quantities are the quadrupole moment and higher moments. However, their expansion is exact only in an infinite complete set; finite expansions will in general lead only to approximations for  $x_{\mu v}$ .

In the following charge and dipole moment analysis of SCF MO calculations, we now use in (2.5) *all* atomic orbitals  $\mu'$  and  $v'$  of the SCF basis set instead of four (or less) orbitals previously [11]. We found the four-orbital expansions rather too short. Their results were often close to Löwdin's two-orbital expansions  $\mu' = \mu$  and  $v' = v$ .

### 3. Calculations and Discussion

We have performed SCF calculations with optimal minimal Slater basis sets expanded in Gaussians on the 4G-level. The program used was MOLPRO written by Meyer and Pulay. The following thirteen molecules were investigated: CO, LiF, LiH, BH, HF, H<sub>2</sub>O, NH<sub>3</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, HCCLi, HCCF. Geometries for the diatomics were taken from Ransil [12], for H<sub>2</sub>O from del Bene and Pople [13], for NH<sub>3</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>2</sub> from Palke and Lipscomb [14] and HCCLi, HCCF from McLean and Yoshimine [15]. In the latter two cases the HC distance chosen was the same as in C<sub>2</sub>H<sub>2</sub>. The exponents for LiH, BH, HF were from Ransil [12], all others from Pople and coworkers [16].

Atomic net charges were calculated for these molecules according to Mulliken, Löwdin and our method in Section 2. These net charges are defined as the differences between gross atomic populations and nuclear charges. The results are presented in Table I. In compounds containing hydrogen, the Mulliken

Table I. Atomic net charges

Molecule	Atom	Mulliken	Löwdin	This work
CO	C	+0.220	+0.245	+0.284
	O	0.220	-0.245	-0.284
LiF	Li	+0.262	+0.340	+0.257
	F	0.262	-0.340	-0.257
LiH	Li	+0.347	+0.618	+0.459
	H	-0.347	-0.618	0.459
BH	B	+0.019	+0.304	+0.153
	H	0.019	0.304	0.153
HF	F	0.234	0.060	0.196
	H	+0.234	+0.060	+0.196
H <sub>2</sub> O	O	0.400	-0.010	-0.300
	H	+0.200	+0.005	+0.150
NH <sub>3</sub>	N	0.492	+0.158	-0.362
	H	+0.164	0.053	+0.121
CH <sub>4</sub>	C	0.080	+0.788	+0.056
	H	+0.020	0.197	-0.014
C <sub>2</sub> H <sub>6</sub>	C	0.045	+0.597	+0.072
	H	+0.015	-0.199	-0.024
C <sub>2</sub> H <sub>4</sub>	C	0.160	+0.276	-0.078
	H	+0.080	-0.138	+0.039
C <sub>2</sub> H <sub>2</sub>	C	-0.198	-0.005	-0.173
	H	+0.198	+0.005	+0.173
HCCLi	H	+0.165	-0.037	+0.132
	C	-0.273	-0.116	-0.293
	Cl	-0.187	-0.248	-0.165
	Li	+0.295	+0.401	+0.326
HCCF	H	+0.207	+0.020	+0.187
	C	-0.217	+0.002	-0.163
	C	+0.082	-0.016	-0.008
	F	-0.072	-0.006	-0.016

analysis leaves the hydrogen more positive by an average amount of 0.035 charge units than our method, except LiH and BH where the relative shift is 0.11 and 0.13. The Löwdin method yields much larger shifts towards hydrogen ranging from 0.17–0.28. The Mulliken analysis disqualifies itself on the basis of non-conservation of the dipole moment of charge distributions. Alternatively, if the two-center atomic dipole moments would be partitioned on the same footing as the charge, the Mulliken analysis would yield atomic hybrid moments on hydrogen, an intolerable situation since there is only a  $1s_H$ -orbital in the basis set. The relative shift of charge towards hydrogen in our method is particularly important in  $CH_4$  and  $C_2H_6$  where the polarity is reversed compared to Mulliken. Since a long time, the organic chemists considered hydrogen as slightly negative in  $CH_4$ . Wheland [17] reasoned that the electronegativity difference of C and H would be overcompensated by the relative effect of size of the C and H atoms in covalent bonds and hybrid moments would also tend to generate a bond dipole moment in which the C is at the positive end. The Löwdin definition is apparently overshooting the corrective effect of charge transfer to hydrogen. It is chemically not appealing to find H to be strongly negative (0.20) in  $CH_4$ , fairly negative (0.05) in  $NH_3$  and almost neutral (0.005) in  $H_2O$ . The explanation is that the Löwdin method corrects the Mulliken method in the right direction, however, its expansion of (2.5) is too short for the stringent condition of dipole moment conservation. The shift of charge towards hydrogen is also strongly pronounced in a method by Politzer and Harris [8]. They applied their method to *standard double-zeta* calculations of  $C_2H_2$ ,  $HCCLi$  and  $HCCF$ . They find a much stronger shift of charge compared to Mulliken for all the atoms than we do with *optimal minimal* basis set calculations. Unfortunately there is no reference to implications for the dipole moment in their work.

For CO and LiF, the differences between Mulliken's, Löwdin's and this work's method are not so pronounced as in most of the other molecules. How relevant the absolute values of these net charges are will be discussed in relation to dipole moments and quality of wavefunctions in one of the following paragraphs.

Table 2 illustrates a breakdown of charge transfer in terms of  $\sigma$  and  $\pi$  separation. From this table it appears that the  $\sigma$  electrons are polarized in the nuclear framework and the  $\pi$  electrons then adjust in the field of nuclei and  $\sigma$  electrons. Thus they are creating an opposite trend compared to the  $\sigma$  electrons. The three types of charge definitions differ only in the magnitude of  $\sigma$  polarization and  $\pi$  adjustment. It should also be pointed out that in minimal basis sets,  $\pi$ -charges are the same in Löwdin's and our method since there is only one  $\pi$  orbital on each atom available for expansion. The second  $\pi$  orbital on any atom cannot be coupled by the dipole operator to any one of the first set.

Table 3 contains the orbital occupation numbers. Most significant for a comparison is the fact that Löwdin's definition strongly depopulates the  $p$ -orbital in the direction of an X–H bond in compounds containing hydrogen when compared with the Mulliken approximation. Our definition reduces this effect considerably. An explanation for the huge transfer by the Löwdin method is given if we consider the center of charge of a  $1s_H 2p_x$  distribution. Since the  $p$ -orbital is directed towards atom H, the center of charge is moved towards H;

Table 2.  $\sigma - \pi$  net charge separation

Molecule	Atom		Mulliken	Löwdin	This work
CO	C	$\sigma$	+0.360	+0.324	+0.363
		$\pi$	-0.140	-0.079	-0.079
	O	$\sigma$	-0.360	-0.324	-0.363
		$\pi$	+0.140	+0.079	+0.079
LiF	Li	$\sigma$	+0.830	+0.800	+0.717
		$\pi$	-0.568	-0.460	-0.460
	F	$\sigma$	-0.830	-0.800	-0.717
		$\pi$	+0.568	+0.460	+0.460
HCCl <sub>3</sub>	H	$\sigma$	+0.165	-0.037	+0.132
		$\pi$	0	0	0
	C	$\sigma$	-0.1235	-0.066	-0.243
		$\pi$	-0.038	-0.050	-0.050
	Cl	$\sigma$	-0.379	-0.410	-0.327
		$\pi$	+0.192	+0.162	+0.162
	Cl	$\sigma$	+0.449	+0.513	+0.438
		$\pi$	0.154	-0.112	-0.112
HCCl	H	$\sigma$	+0.207	+0.020	+0.187
		$\pi$	0	0	0
	C	$\sigma$	0.123	+0.098	-0.067
		$\pi$	0.094	-0.096	-0.096
	Cl	$\sigma$	+0.150	+0.040	+0.048
		$\pi$	-0.068	-0.056	-0.056
	Cl	$\sigma$	-0.234	-0.158	-0.168
		$\pi$	+0.162	+0.152	+0.152

this means a distribution favoring hydrogen. The difference in populations for the three methods is much smaller for the 2s-orbital. Except for the diatomic hydrides and H<sub>2</sub>O both Löwdin's and our definition increase the population of the 2s-orbital compared to Mulliken's and thus generate an opposite effect to the 2p-shift.

Table 4 shows the various parts of the dipole moment in our charge definition. In CO the total hybrid moment is slightly larger than the charge moment and of opposite direction. In essence the atomic moment of the C atom is responsible for the sign of the total dipole moment. The sign of the dipole moment of the wavefunction agrees with the experimental moment [18]. The agreement of magnitude, however, is fortunate since we know that the SCF limit yields the wrong sign [19]. Since the CI results are most often insufficient for good agreement with experiment, it appears at this point rather useless to attempt an analysis with an improved wavefunction. This statement does not hold for LiF. The optimized minimal basis function of LiF yields only 53% of the large experimental dipole moment. Hence, it appears to underestimate the charge moment considerably. Also the relative magnitudes of charge and hybrid moments cannot be considered as final. This also means that the absolute value of net charge in LiF is still greatly underestimated by the best minimal basis set. In all other cases where experimental values are known the wavefunction generates an agreeable dipole moment.

Table 3. Atomic occupation numbers

Molecule	Atom	1s	2s	2p <sub>x</sub>	2p <sub>y</sub>	2p <sub>z</sub>	
CO	C	1.998	1.671	0.570	0.570	0.971	Mulliken Löwdin This work
		1.990	1.714	0.539	0.539	0.972	
		1.987	1.706	0.539	0.539	0.944	
	O	1.999	1.857	1.430	1.430	1.504	
		1.995	1.868	1.461	1.461	1.461	
		1.993	1.883	1.461	1.461	1.486	
LiF	Li	1.991	0.090	0.284	0.284	0.089	
		1.986	0.108	0.230	0.230	0.106	
		1.983	0.167	0.230	0.230	0.132	
	F	1.999	1.940	1.716	1.716	1.891	
		1.997	1.972	1.770	1.770	1.831	
		1.998	1.954	1.770	1.770	1.765	
LiH	Li	1.994	0.402	0	0	0.258	
		1.981	0.293	0	0	0.107	
		1.976	0.392	0	0	0.173	
	H	1.347					
		1.618					
		1.459					
BH	B	1.999	1.809	0	0	1.173	
		1.996	1.811	0	0	0.889	
		1.995	1.807	0	0	1.045	
	H	1.019					
		1.304					
		1.153					
HF	F	1.999	1.946	2	2	1.288	
		1.999	1.934	2	2	1.127	
		1.999	1.918	2	2	1.278	
	H	0.766					
		0.940					
		0.804					
H <sub>2</sub> O	O	1.998	1.862	1.108	2	1.432	
		1.997	1.855	0.887	2	1.270	
		1.997	1.841	1.063	2	1.399	
	H	0.800					
		0.995					
		0.850					
NH <sub>3</sub>	N	1.997	1.597	1.068	1.068	1.762	
		1.988	1.604	0.793	0.793	1.664	
		1.985	1.663	0.990	0.990	1.735	
	H	0.836					
		1.053					
		0.879					
CH <sub>4</sub>	C	1.994	1.141	0.982	0.982	0.982	
		1.973	1.143	0.699	0.699	0.699	
		1.965	1.294	0.895	0.895	0.895	
	H	0.980					
		1.197					
		1.014					

Table 3 (continued)

Molecule	Atom	1s	2s	2p <sub>x</sub>	2p <sub>y</sub>	2p <sub>z</sub>
C <sub>2</sub> H <sub>6</sub>	C	1.994	1.151	0.974	0.974	0.953
		1.974	1.184	0.693	0.693	0.860
		1.967	1.295	0.886	0.886	0.894
	H	0.985				
		1.199				
		1.024				
C <sub>2</sub> H <sub>4</sub>	C	1.995	1.162	1.015	1	0.989
		1.974	1.207	0.704	1	0.838
		1.966	1.314	0.905	1	0.893
	H	0.920				
		1.138				
		0.961				
C <sub>2</sub> H <sub>2</sub>	C	1.997	1.130	1	1	1.071
		1.971	1.197	1	1	0.837
		1.961	1.282	1	1	0.930
	H	0.802				
		0.995				
		0.827				
HCCLi	H	0.835				
		1.037				
		0.868				
	C	1.997	1.161	1.019	1.019	1.076
		1.972	1.243	1.025	1.025	0.852
		1.962	1.329	1.025	1.025	0.953
	C	1.996	1.304	0.904	0.904	1.080
		1.973	1.534	0.919	0.919	0.904
		1.965	1.486	0.919	0.919	0.875
	Li	1.991	0.307	0.077	0.077	0.253
		1.982	0.294	0.056	0.056	0.211
		1.978	0.361	0.056	0.056	0.222
HCCF	H	0.793				
		0.980				
		0.813				
	C	1.996	1.092	1.047	1.047	1.034
		1.971	1.141	1.048	1.048	0.789
		1.960	1.229	1.048	1.048	0.877
	C	1.997	1.010	1.034	1.034	0.844
		1.973	1.166	1.028	1.028	0.822
		1.963	1.192	1.028	1.028	0.797
	F	1.999	1.933	1.919	1.919	1.302
		1.999	1.901	1.924	1.924	1.258
		1.998	1.907	1.924	1.924	1.264

In LiH, FH, NH<sub>3</sub>, HCCLi and HCCF the charge moment dominates greatly over the hybrid moment, in H<sub>2</sub>O they are almost equal and in BH the hybrid moment is considerably larger and of opposite sign to the hybrid moment.



Table 4. Dipole moment partitioning (a.u.)

Molecule	Charge	Hybrid			Total <sup>a</sup>	Exp. <sup>b</sup>
		$X_1$	$X_2$	$X_3$		
CO	-0.605	0.956	-0.312		0.039	0.044
LiF	-0.732	-0.542	-0.101		-1.375	-2.60
LiH	-1.382	-0.941			-2.323	-2.31
BH	-0.357	0.984			0.627	—
FH	0.339	0.233			0.572	0.715
OH <sub>2</sub>	0.381	0.361			0.742	0.706
NH <sub>3</sub>	0.452	0.261			0.713	0.578
HCCLi	1.256	0.029	-0.469	1.075	1.891	.
HCCF	-0.471	0.249	0.221	-0.261	-0.262	-0.279 <sup>c</sup>

<sup>a</sup> Positive entry means - + polarity for the molecule as written.

<sup>b</sup> McClellan, A. L.: Tables of experimental dipole moments. San Francisco: W.H. Freeman and Co. 1963.

<sup>c</sup> Reference from Yoshimine, M., McLean, A. D.: Internat. J. Quantum Chem. 1S, 313 (1967).

Table 5. Partial ionic character of atoms in molecules (%)

Molecule	Electronegativity <sup>a</sup>	Dipole moment <sup>b</sup>	This work <sup>c</sup>
CO	66	2	28
LiF	89	91	26
LiH	26	77	46
BH	< 1	45	15
HF	59	41	20
H <sub>2</sub> O	39	33	15
NH <sub>3</sub>	19	27	12
CH <sub>4</sub>	4		1
C <sub>2</sub> H <sub>6</sub>	4		2
C <sub>2</sub> H <sub>4</sub>	4	---	4
C <sub>2</sub> H <sub>2</sub>	4	-	17

<sup>a</sup> Pauling, L.: The nature of the chemical bond, p. 64, 70. Ithaca, N. J.: Cornell University Press 1948.

<sup>b</sup> Ibid, p. 46.

<sup>c</sup> Charge.

Finally Table 5 summarizes our charge results and compares them with popular definitions by Pauling based on electronegativity differences and experimental dipole moments. We find as Bader and Hanneker [7] did qualitatively that electronegativity is not too reliable to assess the ionic character of an atom in a molecule. Total experimental dipole moments are also limited since we can only estimate or measure by infrared spectroscopy *bond* dipole moments in molecules with vanishing total moments. Bond dipole moments however, are mostly not additive. A weak point in our work is the net charge of LiF which we would expect to be at least as ionic as LiH.

#### 4. Conclusion

The method tested in this paper for atomic charges in molecules is most general and useful in an LCAO SCF approach. It reveals that great differences in charges appear between Mulliken's and Löwdin's definition. The results we obtain by expansions larger than in either of the above methods lie often in between these extremes. To assess the reliability of the suggested method more thoroughly we are in the process of calculating quadrupole moments to see how well the charge definition conserves the quadrupole moment. In principle we know that our method will reproduce a multipole moment with any desired degree of accuracy if we enlarge the basis set for the SCF calculation. This means it is on the same footing as the LCAO approach itself and suffers its drawbacks. An imbalanced basis set might yield a poor charge if the total dipole moment is far off the experimental value. LiF represents such a case. We plan to investigate this molecule with a double zeta set. It should be mentioned also that the convergence of our iterative charge process depends on the main single-center terms of (2.5). Convergence was fast in the hydrides, medium in CO, HCCLi, HCCF and slow in LiF. This seems to indicate somehow the quality of the wavefunction with respect to dipole moments. A question in this context is: Would similar basis sets yield similar charges. This question was asked and partially answered by Politzer and Mulliken [20]. In our definition we would like to tentatively state this: If two wavefunctions yield similar virial quotients and dipole moments and their density distribution in space is similar, we expect similar charges. We shall investigate this point more thoroughly. Finally it should be mentioned that a charge definition based on a least square fit of two-center distributions in single-center distributions is pursued by Meyer [21]. This idea has been used also by Newton [22] and Billingsley and Bloor [23] for integral approximations rather than charges. The method does not conserve the total dipole moment and the charge has to be renormalized. It would be interesting to see how large the difference in charges is between the least square method and our method.

*Acknowledgement.* The main part of this work was done at the Theoretical Chemistry Institute of the University of Stuttgart. I appreciated the hospitality of Prof. H. Preuss and his staff during my stay. Discussions about the topic of this paper with Prof. H. Preuss and Dr. W. Meyer have greatly helped to advance the project. Computer time from the University of Stuttgart and Saint Louis University is gratefully acknowledged. Dr Meyer was helpful in furnishing his program MOLPRO for the SCF calculations.

#### References

1. Mulliken, R.S.: J. Chem. Phys. **23**, 1833 (1955)
2. Daudel, R., Laforge, A.: C.R. seances Acad. Sci. Paris **233**, 623 (1951)
3. Löwdin, P.O.: J. Chem. Phys. **21**, 374 (1953)
4. Christoffersen, R.E., Baker, K.A.: Chem. Phys. Letters **8**, 4 (1971)
5. Löwdin, P.O.: J. Chem. Phys. **18**, 365 (1950)
6. Pople, J.A., Beveridge, D.L.: Approximate molecular orbital theory. New York: McGraw-Hill 1970
7. Bader, R.F.W., Hanneker, W.H.: J. Am. Chem. Soc. **88**, 280 (1966)
8. Politzer, P., Harris, R.R.: J. Am. Chem. Soc. **92**, 6451 (1970)

9. Hartmann, H., Jug, K.: *Theoret. Chim. Acta (Berl.)* **3** (1965)
10. Jug, K.: *Theoret. Chim. Acta (Berl.)* **26**, 231 (1972)
11. Jug, K.: *Theoret. Chim. Acta (Berl.)* **29**, 9 (1973)
12. Ransil, B.: *Rev. Mod. Phys.* **32**, 245 (1960)
13. del Bene, J., Pople, J. A.: *Chem. Phys. Letters* **4**, 426 (1969)
14. Palke, W. E., Lipscomb, W.: *J. Am. Chem. Soc.* **88**, 2384 (1966)
15. McLean, A. D., Yoshimine, M.: *Tables of linear molecular wavefunctions*, IBM 1967
16. Hehre, W. J., Stewart, R. F., Pople, J. A.: *J. Chem. Phys.* **51**, 2657 (1969)
17. Wheland, G. W.: *Resonance in organic chemistry*, p. 204ff. New York: Wiley 1955
18. McLean, A. D., Yoshimine, M.: *J. Chem. Phys.* **45**, 3467 (1966)
19. Schaefer, III, H. F.: *The electronic structure of atoms and molecules*, p. 196ff. Reading: Addison-Wesley 1972
20. Politzer, P., Mulliken, R. S.: *J. Chem. Phys.* **55**, 5135 (1971)
21. Meyer, W.: Private communication
22. Newton, M. D.: *J. Chem. Phys.* **51**, 3917 (1969)
23. Billingsley, F. P., Bloor, J. E.: *J. Chem. Phys.* **55**, 5178 (1971)

Prof. Dr. K. Jug  
Department of Chemistry  
Saint Louis University  
Saint Louis, Missouri 63156, USA



## *Ab initio* Calculations on Large Molecules Using Molecular Fragments. Characterization of the Zwitterion of Glycine\*

Lester L. Shipman\*\* and Ralph E. Christoffersen\*\*\*

Department of Chemistry, University of Kansas, Lawrence, Kansas 66044

Received April 19, 1973

The *ab initio* molecular fragment approach is applied to a characterization study of the ground state of the zwitterion of glycine. Included among the properties studied are the  $\phi - \psi$  conformational energy surface, the electronic structure, and the magnitude and direction of the dipole moment. The results of the present study are compared to the results of other theoretical and experimental studies.

**Key words:** Zwitterion of glycine Conformational energy surface – *Ab initio* molecular fragment approach.

### 1. Introduction

Glycine, a biologically important molecule having the distinction of being the smallest amino acid, exists as a zwitterion (see Fig. 1) in aqueous solution in the range of biological pH's [1]. In previous theoretical studies, the  $\phi - \psi$  conformational energy surface of the zwitterion of glycine has been studied by Ponnuswamy and Sasisekharan [2] using an empirical potential energy function; Imamura *et al.* [3a] and Oegerle *et al.* [3b] have studied the electronic structure and preferred geometry of the zwitterion and neutral forms of glycine in CNDO studies; and the zwitterion of glycine has been characterized in the X-ray crystallographic geometry [4, 5] of  $\alpha$ -glycine in an *ab initio* study by Ryan and Whitten [6].

In the present study, the *ab initio* SCF molecular fragment approach [7–19] is applied to a characterization study of the ground state of the zwitterion of glycine. These studies serve partly to provide comparisons with experimental and other theoretical studies, and partly to elucidate the ability of the molecular fragment approach to describe systems with large charge polarization. The entire  $\phi - \psi$  conformational energy surface is studied, and a  $\phi - \psi$  conformational energy contour map is presented. The molecular orbital structure, as well as the magnitude and direction of the dipole moment vector are discussed for the conformation of lowest energy.

\* This work was supported in part by the National Science Foundation, the University of Kansas, and the Upjohn Company, Kalamazoo, Michigan 49001.

\*\* NSF Trainee (1969–1972). Current address — Department of Chemistry, Cornell University, Ithaca, New York 14850.

\*\*\* Alfred P. Sloan Research Fellow (1971–1973). Author to whom correspondence should be directed.

## ZWITTERION OF GLYCINE

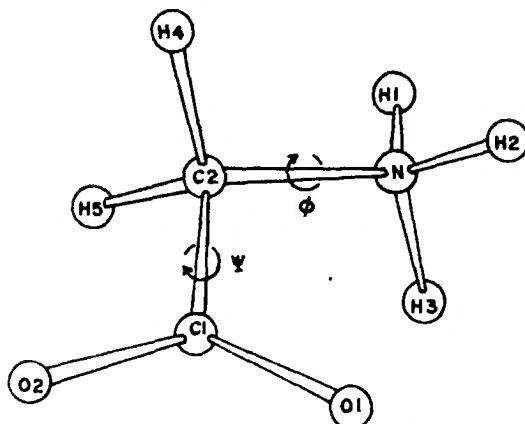


Fig. 1. Nuclear numbering system and definition of  $\phi$  and  $\psi$  for the zwitterion of glycine

The nuclear geometry used in the present study is taken from the recently determined neutron diffraction geometry of  $\alpha$ -glycine by Jönsson and Kvik [20]. This geometry is modified slightly according to several constraints, so that many fewer calculations are needed to construct a grid for the entire conformational energy surface in  $\phi$  ( $\angle$  H3-N-C2-C1) and  $\psi$  ( $\angle$  N-C2-C1-O1) [21]. The convention that  $(\phi, \psi) = (0, 0)$  [22] is the fully eclipsed conformation is used, and the constraints applied are the following:

- 1) A  $\phi$ -rotation of  $120^\circ$  results in a final structure that is identical to the initial structure.
- 2) A  $\psi$ -rotation of  $180^\circ$  results in a final structure that is identical to the initial structure.
- 3) Reflection through the C1-C2-N plane takes H4 identically into H5, and vice versa.

The imposition of these constraints results in only small deviations from the experimental structure (see Table I). The largest bond distance deviation is  $0.015 \text{ \AA}$  and the RMS deviation is  $0.007 \text{ \AA}$ , while the largest bond angle deviation is  $1.2^\circ$ , and the RMS deviation is  $0.6^\circ$ . With these three constraints imposed, the entire conformational energy surface may be studied with  $\phi$  running from  $0$  to  $120^\circ$  and  $\psi$  running from  $0$ – $180^\circ$  and also, the surface has inversion symmetry about the point  $(60, 90)$ .

The molecular fragment basis set [7–19], consisting of 30 floating spherical Gaussian functions (FSGO) [23] that are contracted to 25 basis functions, is described in Table 2. Since the methods for determining these basis orbitals and for forming large molecules from them have been described earlier, [7–19] they will not be repeated here.

## 2. $\phi$ – $\psi$ Conformational Energy Surface

A grid over the conformational energy surface [24] was calculated in  $15^\circ$  increments in  $\phi$  and  $\psi$ . The absolute minimum was found at the point  $(60, 0)$ . Using linear interpolation, a contour map was constructed (see Fig. 2) with

Table 1. Nuclear geometry

Bond	Experimental <sup>a</sup>	Constrained <sup>b</sup>
Bond distances (in Å)		
C1-O1	1.250	1.250
C1-O2	1.251	1.250
C1-C2	1.526	1.526
C2-H4	1.090	1.090
C2-H5	1.089	1.090
C2-N	1.476	1.476
N-H1	1.054	1.039
N-H2	1.037	1.039
N-H3	1.025	1.039
Bond angles (in degrees)		
O2-C1-O1	125.4	125.4
O1-C1-C2	117.5	117.3
O2-C1-C2	117.1	117.3
C1-C2-N	111.9	111.9
C1-C2-H4	108.8	109.7
C1-C2-H5	110.5	109.7
H4-C2-H5	108.0	108.0
N-C2-H4	108.5	108.8
N-C2-H5	109.0	108.8
C2-N-H1	112.1	111.4
C2-N-H2	111.7	111.4
C2-N-H3	110.4	111.4
H1-N-H2	108.7	107.5
H1-N-H3	107.1	107.5
H2-N-H3	106.6	107.5

<sup>a</sup> See Ref. [20].<sup>b</sup> The geometry used in the present studies.Table 2. Molecular fragment data<sup>a</sup>

Fragment type	FSGO type	FSGO Distance from "heavy" atom	FSGO radii (ρ)
CH <sub>4</sub> (T <sub>d</sub> )	C-H	1.23379402	1.67251562
R(C, H) = 2.05982176	C inner shell	0.0	0.32784375
CH <sub>3</sub> (planar)	C-H	1.13093139	1.51399487
R(C, H) = 1.78562447	C-π	±0.1	1.80394801
	C inner shell	0.0	0.32682735
OH(sp hybrid)	O-H	0.76467773	1.23671871
R(O, H) = 1.54774058	O-LP(σ)	0.21614258	1.28753780
	O-LP(P)	±0.1	1.19741696
	O-π	±0.1	1.12242182
	O inner shell	0.00057129 <sup>b</sup>	0.24028227
NH <sub>4</sub> <sup>+</sup> (T <sub>d</sub> )	N-H	0.80547793	1.50046875
R(N, H) = 1.95021656	N inner shell	0.0	0.27770068

<sup>a</sup> All distances are in Hartree atomic units, see Shull, H., Hall, G. G.: Nature **184**, 1559 (1959).<sup>b</sup> This is the distance from the oxygen nucleus, along the O-H bond axis, toward the H nucleus.

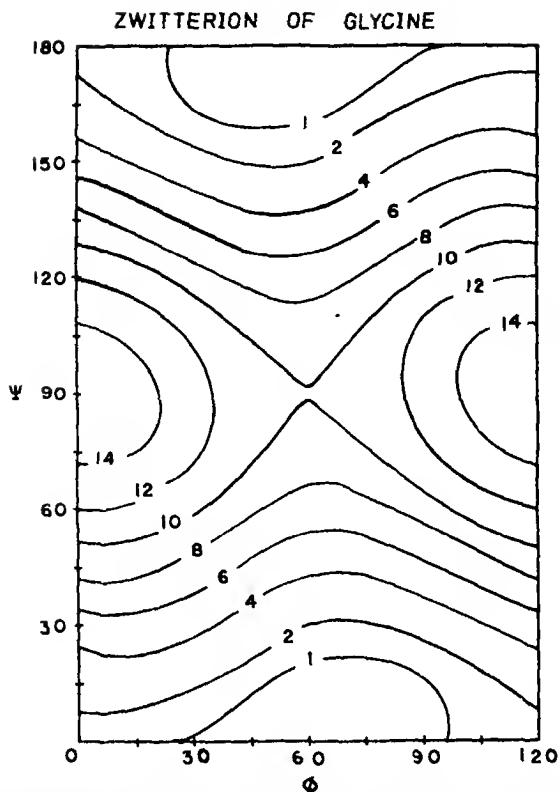


Fig. 2  $\phi - \psi$  conformational energy surface for the zwitterion of glycine. Contours are given in units of kcal/mole relative to the energy of the conformation of lowest energy

contours in kcal/mole relative to the absolute minimum. Note that the conformational energy surface is saddle-shaped with a saddle point at  $(60, 90)$ . Also, the  $\phi - \psi$  map indicates that a wide range of conformations is accessible. For example, within a 2 kcal/mole range, any value of  $\phi$  is accessible when  $\psi = 0^\circ$ . For a rotation starting and terminating at the absolute minimum energy conformation, the lowest barrier for  $\phi$ -rotation is 1.5 kcal/mole at  $(0, 0)$  and the lowest barrier for  $\psi$ -rotation is 10.1 kcal/mole at  $(60, 90)$ .

Using an empirical energy function, Ponnuswamy and Sasisekharan [2] have also found the minimum to be at  $(60, 0)$ , and 0.5 and 1 kcal/mole contours were plotted. It is of interest to compare the extent from the minimum of the 1 kcal/mole contour in that study and in the present study. This contour has roughly the same maximum extent in the  $\phi$ -direction in the two studies, with the contour of the present study having the largest extent. The greatest extent of the 1 kcal/mole contour in the  $\psi$ -direction in the study of Ponnuswamy and Sasisekharan is to  $\psi \approx 60^\circ$ , while in the present study this contour extends to  $\psi \approx 20^\circ$ .

Imamura *et al.* [3a], in a CNDO study, fixed  $\phi$  at  $60^\circ$  and found that  $\psi = 0^\circ$  was the energetically preferred conformation. They found a barrier for  $\psi$ -rotation of 5.8 kcal/mole at  $(60, 90)$ .



There are three crystalline forms of glycine,  $\alpha$  [4, 5, 20],  $\beta$  [25], and  $\gamma$  [26], which differ primarily in the network of hydrogen bonds between zwitterions.  $\phi$  is known to be roughly  $60^\circ$  in  $\alpha$ -glycine [20] but  $\phi$  has not been determined in  $\beta$  and  $\gamma$ -glycine.  $\psi$  is known to be approximately  $18.6^\circ$  in  $\alpha$ -glycine,  $24.8^\circ$  in  $\beta$ -glycine, and  $12.8^\circ$  in  $\gamma$ -glycine [26]. Note that, for a range of values of  $\phi$  (see Fig. 2), these  $\psi$  values correspond to conformations that are less than 2 kcal/mole from the minimum energy conformation. In particular (58.3, 18.6), is the point in the constrained geometry closest to the experimental geometry of  $\alpha$ -glycine [20], and this point has an energy 0.8 kcal/mole above the minimum. However, comparison of the preferred conformation predicted by the present study with the conformation of the monomer in  $\alpha$ -glycine from neutron diffraction must be made with caution. The results of the present study apply to an isolated molecule, and the effects of the crystal environment on monomer conformation could be considerable in the case of  $\alpha$ -glycine. For example, Oegerle and Sabin [3b] have estimated that the displacement of the nitrogen atom from the plane of the heavy atoms in the crystal is a crystal packing effect, that can occur with only very little cost in energy of the zwitterion. Thus, the preferred values of  $\phi$  and  $\psi$  may be determined by orientational requirements of *intermolecular* hydrogen bonding in the crystal, since O1, O2, H1, H2, H3, (and possibly H4 and H5) may participate in intermolecular hydrogen bonds. However, the high monomer conformational energy associated with values of  $\psi$  near  $90^\circ$  should be sufficient to make these values unfavorable in the crystal.

### 3. Dipole Moment

The dipole moment,  $\mu$ , of the zwitterion of glycine has been determined in the present study in the minimum energy conformation, (60, 0).  $\mu$  is constrained by symmetry to lie in the C1-C2-N plane.  $\mu$  has a magnitude of 13.33 D [27] and forms an angle of  $29.1^\circ$  with the C1-C2 bond. This result is in excellent agreement with experimental value of 13.3 D for the zwitterion of glycine in aqueous solution as measured by Buckingham [28], and in moderate agreement (11%) with the experimental value of 15.0 D for the zwitterion of glycine in water-alcohol solution as measured earlier by Kirkwood [29]. Ryan and Whitten [6], in an *ab initio* study, have calculated the dipole moment of the zwitterion of glycine in the X-ray crystallographic geometry to be 12.17 D, while Oegerle and Sabin [3b] have calculated a value of 13.410 D for a geometry optimized using the CNDO procedure.

### 4. Electronic Structure

The electronic structure of the zwitterion of glycine has been studied at the minimum energy conformation, (60, 0), and at the point (in the constrained geometry) closest to the neutron diffraction geometry [20], (58.3, 18.6). These results are presented in Table 3, along with the results from the *ab initio* study of Ryan and Whitten [6] at the X-ray crystallographic geometry [4, 5]. In the present study, each molecular orbital is characterized by means of a population analysis over the symmetric orthonormalized basis functions [18]. As has been generally observed in previous studies [7-19], the ordering of the valence molecular orbitals is identical to that observed in more extensive basis set studies. Further-

Table 3. Electronic structure<sup>a</sup>

(φ, ψ) Energy minimum <sup>b</sup>		Experimental geometry <sup>c</sup>		Ryan and Whitten <sup>f</sup>	
Present calc.		Present calc. <sup>e</sup>			
MO type <sup>d</sup>	ε <sub>i</sub>	MO type <sup>d</sup>	ε <sub>i</sub>	MO type <sup>d</sup>	ε <sub>i</sub>
O 2 (1s)	-16.893	O 2 (1s)	-16.893	O 2 (1s)	-20.363
O 1 (1s)	-16.865	O 1 (1s)	-16.865	O 1 (1s)	-20.348
N (1s)	-13.219	N (1s)	-13.220	N (1s)	-15.885
C 2 (1s)	-9.318	C 2 (1s)	-9.318	C 1 (1s)	-11.488
C 1 (1s)	-9.126	C 1 (1s)	-9.126	C 2 (1s)	-11.448
N H	-1.322	N H	-1.323	N-H	-1.500
C-O(σ)	-1.127	C-O(σ)	-1.127	C-O(σ)	-1.344
C-O(σ)	1.022	C-O(σ)	-1.021	C-O(σ)	-1.221
C 11, C' C'	-0.894	C H, C' C'	-0.895	σ	-1.123
N H	-0.664	N H	-0.665	N-H	-0.963
N H	-0.659	N H	-0.659	N-H	-0.946
C N	0.539	C N	-0.540	σ	-0.807
C 11, C' C'	0.477	C H, C' C'	-0.477	σ	-0.751
C H	0.452	C H	-0.451	σ	-0.736
O(sp - lp) <sub>1</sub>	-0.250	O(sp - lp) <sub>1</sub>	-0.249	O 1, O 2 (lp)	-0.575
C-O(σ)		C-O(σ)			
π <sub>b</sub>	-0.217	π <sub>b</sub>	-0.217	π <sub>b</sub>	-0.523
O(sp - lp)	0.160	O(sp - lp)	-0.159	O 1, O 2 (lp)	-0.504
O(P - lp)	+0.025	O(p - lp)	+0.026	O 1, O 2 (lp)	-0.358
O(P - lp)	+0.037	O(P - lp)	+0.038	O 1, O 2 (lp)	-0.339
π <sub>a</sub>	+0.087	π <sub>a</sub>	+0.088	π <sub>a</sub>	-0.289
π <sub>a</sub>	+0.680 <sup>g</sup>	π <sub>a</sub>	+0.654 <sup>g</sup>		

<sup>a</sup> All energies are reported in Hartree atomic units. See Shull, H., Hall, G.G.: *Nature* **184**, 1559 (1959).

<sup>b</sup> (φ, ψ) = (60, 0)

<sup>c</sup> The neutron diffraction geometry (Ref. [20]), X-ray crystallographic geometry (Refs. [4] and [5]), and constrained geometry are quite similar. See the text for a discussion of the differences between the constrained and neutron diffraction geometry. The X-ray geometry is essentially the same as the neutron diffraction geometry except that the heavy atom-hydrogen bond distances are 13-17% smaller.

<sup>d</sup> Where a given molecular orbital has a contribution of at least 50% (1.0 electron) of a particular type (as measured by a population analysis over the symmetric orthonormalized basis functions), then that contribution is indicated symbolically. If no single contribution comprises 50% of the molecular orbital, then the primary and secondary contributions are listed, separated by a comma.

<sup>e</sup> (φ, ψ) = (58.3, 18.6)

<sup>f</sup> These results are from an *ab initio* SCF calculation of Ryan and Whitten (Ref. [6]) at the X-ray crystallographic geometry, using a basis set of 152 spherical Gaussian functions. The notation used to label the molecular orbitals has been changed from that found in Ref. [6] to an equivalent notation, in some instances.

<sup>g</sup> Lowest unoccupied molecular orbital.

more, the detailed nature of the molecular orbitals is seen to be identical at the level of characterization in the Ryan and Whitten study, except that the ordering of the inner shell molecular orbitals involving C1 and C2 is reversed. Disappointingly, the highest three occupied molecular orbitals have positive orbital energies in the present study. However, a plot of the valence molecular orbital energies, ε<sub>i</sub> (R-W), for the Ryan and Whitten study vs. the corresponding orbital

energies,  $\epsilon_i$ , for the present study at (58.3, 18.6) is approximately linear. This type of behavior has been observed in a wide variety of examples using molecular fragment basis sets [30], and appears to be an indication of the good balance of basis sets obtained using the molecular fragment procedure. A least squares fit of the current data to a straight line results in the following equation,

$$\epsilon_i(\text{R-W}) = 0.8571 \cdot \epsilon_i - 0.362, \quad (1)$$

where the RMS deviation from this straight line is 0.016. Thus, even though the molecular orbital energies have been shifted upward and the spacings increased slightly, the proper ordering of valence molecular orbitals is maintained. Note also that the character of the various molecular orbitals is the same for the (60, 0) and (58.3, 18.6) conformations, and that the molecular orbital energies are nearly the same, with the largest change occurring for the orbital energy of the lowest unoccupied molecular orbital.

### 5. Discussion and Summary

One of the more striking aspects of this study is the appearance of a rather wide range of conformations that are accessible to the isolated zwitterion, that are in the vicinity of the lowest energy conformer (60, 0). For example, the expected ethane-type rotational barrier for rotation about  $\phi$  (for  $\psi = 0^\circ$ ) is substantially lower (1.5 kcal/mole) than is calculated in ethane itself (5.60 kcal/mole [8]) using the molecular fragment procedure. This rather flat energy surface for  $\phi$ -rotation is apparently due to the presence of an additional effect in the case of the glycine zwitterion that is not present in ethane itself. In particular, the nuclear positions for an optimum intramolecular hydrogen bond between H3 and O1 are present at approximately (0, 0). Thus, the stabilization of the molecules by an intramolecular hydrogen bond offsets the destabilization that is due to an "ethane-type" eclipsing at (0, 0), and results in a rather flat ( $\phi$ , 0) curve.

Because of the low conformational energies associated with a  $\phi$ -rotation and the fact that the intramolecular hydrogen bonded conformation, (0, 0), lies only 1.5 kcal/mole above the lowest calculated energy conformation, the zwitterion of glycine probably assumes those values of  $\psi$  in aqueous solution and in the crystalline state that facilitate intermolecular hydrogen bonding to the solvent water molecules in the former case and to other zwitterions in the latter case. Because of the higher energies associated with  $\psi$ -rotations, it is anticipated that the values of  $\psi$  will remain in the  $0^\circ \leq \psi \leq 30^\circ$  range. However, within this range, those values of  $\psi$  which facilitate intermolecular hydrogen bonding in aqueous solution and in the crystalline state would be expected to be favored, as in the case of  $\phi$ .

The dipole moment,  $\mu$ , was found to have a magnitude of 13.33 D, in good agreement with the experimental values, and  $\mu$  was found to form an angle of 29.1 with the C1-C2 bond. The molecular orbital ordering obtained in the present study was in agreement with the ordering obtained by Ryan and Whitten [6], and an approximately linear relationship was found to hold between the valence molecular orbital energies calculated by Ryan and Whitten and those calculated in the present study.

Thus, even for molecules in which rather large charge polarization is encountered, the molecular fragment procedure appears to provide a well balanced description of several molecular properties of interest, and appears to be appropriate for description of molecular systems of considerable size and variety.

*Acknowledgement.* The authors would like to express their appreciation to the University of Kansas for partial support of the computing costs required for these studies.

### References

- Hückel, W.: Theoretical principles of organic chemistry, Vol. II, 150-157. New York: Elsevier 1958
- Ponnuswamy, P. K., Sasisekharan, V.: Biopolymers 7, 624 (1969)
- Imamura, A., Fujita, H., Nagata, C.: Bull. Chem. Soc. Japan 42, 3118 (1969)
- Oegerle, W. R., Sabin, J. R.: J. Molec. Struct. 15, 131 (1973)
- Albrecht, G., Corey, R. B.: J. Am. Chem. Soc. 61, 1087 (1939)
- Marsh, R. E.: Acta Cryst. 11, 654 (1958)
- Ryan, J. A., Whitten, J. L.: J. Am. Chem. Soc. 94, 2396 (1972)
- Christoffersen, R. E., Maggiora, G. M.: Chem. Phys. Lett. 3, 419 (1969)
- Christoffersen, R. E., Genson, D. W., Maggiora, G. M.: J. Chem. Phys. 54, 239 (1971)
- Christoffersen, R. E.: J. Am. Chem. Soc. 93, 4104 (1971)
- Christoffersen, R. E.: Adv. Quan. Chem. 6, 333 (1972)
- Maggiora, G. M., Genson, D. W., Christoffersen, R. E., Cheney, B. V.: Theoret. chim. Acta (Berl.) 22, 337 (1971)
- Christoffersen, R. E., Shipman, L. L., Maggiora, G. M.: Int. J. Quant. Chem. 5S, 143 (1971)
- Cheney, B. V., Christoffersen, R. E.: J. Chem. Phys. 56, 3503 (1972)
- Genson, D. W., Christoffersen, R. E.: J. Am. Chem. Soc. 94, 6904 (1972)
- Shipman, L. L., Christoffersen, R. E.: Chem. Phys. Lett. 15, 469 (1972)
- Genson, D. W., Christoffersen, R. E.: J. Am. Chem. Soc. 95, 362 (1973)
- Shipman, L. L., Christoffersen, R. E.: Proc. Nat. Acad. Sci. US 69, 3301 (1972)
- Shipman, L. L., Christoffersen, R. E.: J. Am. Chem. Soc. 95, 1408 (1973)
- Weiman, L. J., Christoffersen, R. E.: J. Am. Chem. Soc. 95, 2074 (1973)
- Jönsson, P.-G., Kvick, Å.: Acta Cryst. B28, 1827 (1972)
- For a  $15^\circ$  grid in  $\phi$  and  $\psi$ , 576 calculations would be required for the experimental geometry, while only 50 calculations are required for the constrained geometry
- Throughout this manuscript conformations will be specified by the notation,  $(\phi, \psi)$
- These basis orbitals were introduced and utilized for small molecules by Frost, A. A., and co-workers: J. Chem. Phys. 54, 764 (1971), and in earlier references contained therein. Formulation and characterization of these orbitals for large molecule calculations is given in [7-19]
- Since the  $\phi - \psi$  surface has inversion symmetry about the point  $(60, 90)$ , the discussion of this surface will be limited to the region  $0^\circ \leq \phi < 120^\circ$ ,  $0^\circ \leq \psi \leq 90^\circ$ , without loss of generality
- Itaka, Y.: Acta Cryst. 13, 35 (1960)
- Itaka, Y.: Acta Cryst. 14, 1 (1961)
- The dipole moment has been calculated according to the conversion factor, 2.54158 D/a.u.
- Buckingham, A. D.: Aust. J. Chem. 6, 323 (1953)
- Kirkwood, J. G.: J. Chem. Phys. 2, 351 (1934)
- This approximately linear relationship between orbital energies from extensive basis set calculations and those of the molecular fragment procedure has been observed in 30 molecules to date, and is apparently not restricted to any specific type of molecule. For example, the cases studied to date include linear and non-linear molecules, saturated and unsaturated molecules, and molecules with or without hetero-atoms

Prof. Dr. R. E. Christoffersen  
Department of Chemistry  
University of Kansas  
Lawrence, Kansas 66044  
USA

## An *ab initio* Study of the $A_{Ac}1$ Hydrolysis Mechanism of Formamide

A. C. Hopkinson

Department of Chemistry, York University, Downsview, Ontario, Canada

I. G. Csizmadia

Lash Miller Chemical Laboratories, University of Toronto, Toronto 181, Ontario, Canada

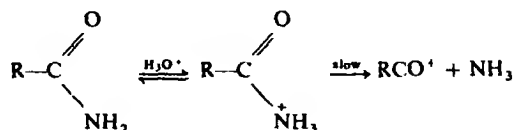
Received March 26, 1973

*Ab initio* calculations have been used to examine the reaction profile for the  $A_{wc}1$  hydrolysis mechanism for formamide, giving a value of 67.3 kcal/mole for  $\Delta H^\ddagger$ . Comparisons between computed and experimental proton affinities are used to assess the reliability of the calculations. Orbital energies are reported for formamide, N-protonated and O-protonated formamide, carbon monoxide and the formyl cation.

**Key words:** Protonation of Formamide — Geometry of Formyl Cation — Reaction Profile — Computed Activation Energy.

We have previously reported a theoretical study [1] on the acid catalysed hydrolysis of amides, esters and carboxylic acids ( $O^{18}$  exchange). The results of a small basis set *ab initio* calculation on the formic acid "hydrolysis" were markedly different from the majority of the calculations, which employed the semi-empirical extended Hückel molecular orbital (EHMO) method. Recent *ab initio* studies [2, 3] have also shown that the EHMO predictions on the site of protonation of formamide were also incorrect. It was therefore decided to reexamine the  $A_{wc}1$  reaction profile for the decomposition of protonated formamide, this time using the more reliable *ab initio* method.

Experimentally the reaction, which occurs in concentrated mineral acid solutions, is believed to proceed through a fast protonation step followed by a slow decomposition of the nitrogen protonated tautomer to produce the acylium ion and ammonia.



Acylium ions are stable only as salts with  $\text{SbF}_6^-$  ions [4–6] or in oleum solutions [7], and quickly pick up water in most acidic solutions to produce carboxylic acids. In the present theoretical study only formamide ( $\text{R}=\text{H}$ ) was considered.

### Computational Details

The calculations were all carried out on an IBM 360-65, using a modified version of IBMOL-IV [8]. The basis set consisted of 8<sup>s</sup> and 3<sup>p</sup> type Gaussian functions on oxygen, nitrogen and carbon, and 3<sup>s</sup> type functions on hydrogen, contracted to 2<sup>s</sup>, 1<sup>p</sup> and 1<sup>s</sup> respectively [9].

The molecular geometries were those used in previous theoretical treatments [10-12]. Previous workers [13] using minimal basis set calculations found the carbon-oxygen bond length of  $\text{HCO}^+$  to be 0.02 a.u. *shorter* than in carbon monoxide. In the initial calculations in the present study the geometry of this ion was reoptimised with the larger basis set and was again found to be linear, but with the carbon-oxygen bond 0.075 a.u. *larger* than in carbon monoxide [14] (Fig. 1).

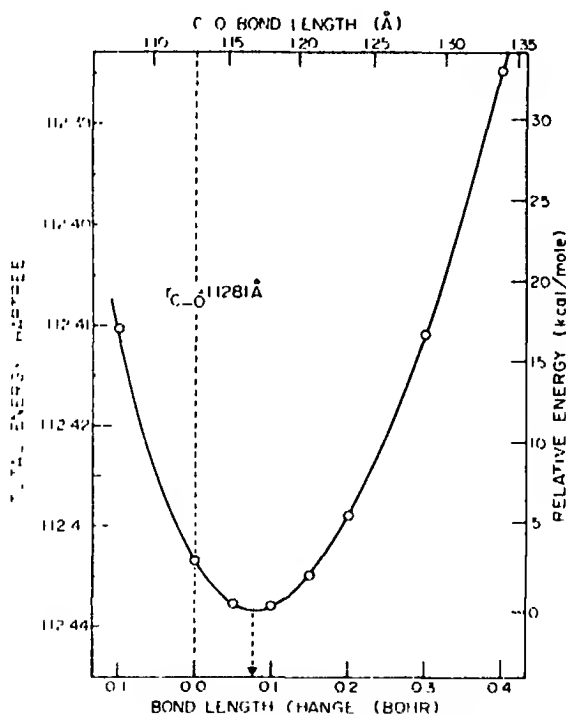


Fig. 1 Plot of computed total energy against carbon-oxygen bond length for the formyl ion

### Results and Discussion

The computed total energies for all the species are recorded in Table 1 along with the estimated Hartree Fock limits [15, 12, 11]. These were used to construct the reaction profile shown in Fig. 2. The *O-protonated* formamide, although not a possible intermediate in this mechanism, was included to show the difference between the N- and O-*protonated* tautomers. The initial protonation reaction is

Table 1. Computed total energies (hartrees)

Species	Calculated total energy	Hartree Fock limit
	- 168.1296 (planar) - 168.1299 (experimental geometry)	- 168.942 <sup>a</sup>
	- 168.5358	—
	- 168.5259	—
	- 187.8161	- 188.957 <sup>b</sup>
	- 188.1576	—
H <sub>2</sub> O	- 75.6121	- 76.048 <sup>c</sup>
H <sub>3</sub> O <sup>+</sup>	- 75.9726	—
NH <sub>3</sub>	- 55.9452	- 56.190 <sup>c</sup>
NH <sub>4</sub> <sup>+</sup>	- 56.3512	—
C O	- 112.2118	- 112.781 <sup>c</sup>
HCO <sup>+</sup>	- 112.4380	—

<sup>a</sup> See Ref. [15].<sup>b</sup> See Ref. [12].<sup>c</sup> See Ref. [11].

predicted to be exothermic by 22.3 kcal/mole. However this energy difference, neglecting changes in zero point vibrational energy, is  $\Delta H_0^\circ$ , the heat of reaction at 0°K. The equilibrium at 25°C depends on  $\Delta G_{298}^\circ$ , a thermodynamic property which has large contributions from enthalpies and entropies of solution. The hydration energy of the hydronium ion (H<sub>3</sub>O<sup>+</sup>) alone is estimated to be

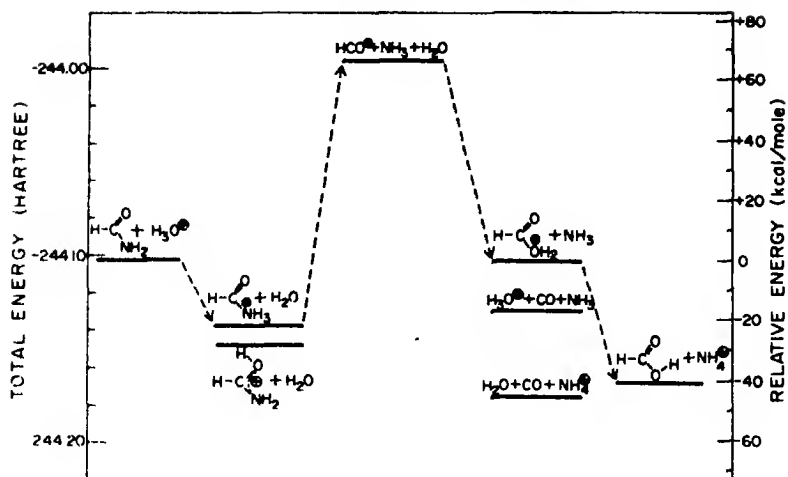
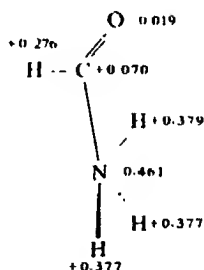


Fig. 2. Computed reaction profile for the  $A_{A1}$  hydrolysis mechanism for formamide

100 kcal/mole, [15] and this ion due to its ability to form strong hydrogen bonds will probably be the species most stabilised by solvation. Solvation then could easily reverse the relative energies of the reactants and products of the first step on the reaction profile.

Experimentally the conversion of the *N*-protonated amide to the acylium ion and ammonia is the slow step. Formally this requires heterolytic fission of the carbon-nitrogen bond and will be assisted by the partial positive charge (+0.67) on the  $NH_3$  group.



The products of this fission reaction are computed to be 89.6 kcal/mole less stable than the *N*-protonated formamide. If these unstable products have similar energy to the transition state for this reaction, [17] then  $\Delta H^*$  for the hydrolysis is calculated to be 67.3 kcal/mole. This compares with experimental values of between 27.2 and 31.8 kcal/mole for different benzamides [18].

Decomposition of the formyl ion into carbon monoxide is computed to be more favourable than nucleophilic attack by water to produce formic acid. All these processes are predicted to be very exothermic, consistent with the experi-



mental observation that the formyl ion is very unstable [19]. The most stable products are carbon monoxide, water and the ammonium ion, the usual products found in the decomposition of formamide.

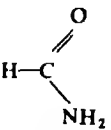
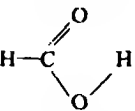
The overall profile is similar to that obtained in *ab initio* studies on the  $A_{Ac}1$  decomposition of formic acid [1, 12], and is markedly different from that produced by the EHMO calculations. In particular, it predicts high energy intermediates as products of the kinetically slow step, whereas the semi-empirical calculations gave these to be the most stable species on the profile.

In an attempt to assess the accuracy of calculations with this particular basis set, the proton affinities for the several weak bases were compared with experimental values. These are recorded in Table 2.

*Hydroxy-protonated* formic acid, the only tautomer used in construction of this profile, was previously found to be 25 kcal/mole less stable than the *cis-trans* carbonyl protonated formic acid [12]. The computed proton affinity for formic acid in Table 2 will therefore be too high by this amount. All the computed proton affinities overestimate the experimental values, but the energy differences vary considerably. The largest difference is for water, for which the calculation overestimates the experimental value by almost 50 %. Therefore it is safe to conclude that errors due to inadequacies in the basis set lead to an overestimation of energy differences for *gas phase* reactions of at most a factor of two.

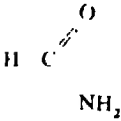
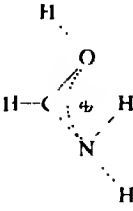
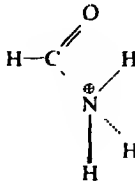
Finally, the orbital energies of protonated formamide and the formyl cation have never been reported. These are listed in Table 3, along with those of the unprotonated bases. Similarly to other species [21], protonation clearly stabilises *all* molecular orbitals, including those composed mainly of core orbitals from the carbon, nitrogen and oxygen atoms. This analysis of orbital energies then clearly

Table 2. Proton affinities (kcal/mole)

Molecule	Computed value	Experimental value
H <sub>2</sub> O	226.3	151 <sup>a</sup>
NH <sub>3</sub>	254.9	200 <sup>a</sup>
CO	142.0	131 <sup>a</sup>
 (O-protonation)	254.8	—
 (hydroxy-protonation)	214.4	162 <sup>b</sup>

<sup>a</sup> See Ref. [11].<sup>b</sup> See Ref. [20].

Table 3. Computed orbital energies

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>H-C-NH<sub>2</sub></p> </div> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> <div style="text-align: center;"> <p>CO</p> </div> <div style="text-align: center;"> <p>HCO<sup>+</sup></p> </div> </div>				
20.6062	20.9956	- 20.9908	- 20.8203	- 21.3169
15.6202	15.9978	- 16.0895	- 11.4371	- 11.9789
11.4492	11.8096	11.7342	- 1.5294	- 1.8920
1.3719	- 1.6755	- 1.6428	- 0.7529	- 1.2155
1.1527	- 1.4816	- 1.5373	- 0.6071	- 1.0618
- 0.8006	- 1.1419	- 1.0949	- 0.6071	- 0.9721
0.6941	- 1.0311	- 1.0109 (a'')	- 0.5001	- 0.9721
0.6165	- 0.9434	- 1.0090		
0.5361	0.9084	- 0.9277		
0.5295	0.8340 (a'')	0.8285		
0.3622	0.7543	- 0.7243 (a'')		
0.3362	0.6700 (a'')	- 0.6784		

shows that valence shell molecular orbital calculations (like the EHMO method), which assume the energies of the inner shell electrons are unchanged, cannot hope to successfully predict energy differences for even the simplest chemical reaction, protonation.

*Acknowledgments* The authors wish to thank the National Research Council for continued financial support.

### References

1. Hopkinson, A.C., McClelland, R.A., Yates, K., Csizmadia, I.G.: *Theoret. Chim. Acta (Berl.)* **13**, 65 (1969)
2. Bonaccorsi, R., Pullman, A., Scrocco, E., Tomasi, J.: *Chem. Phys. Letters* **12**, 622 (1972)
3. Hopkinson, A.C., Csizmadia, I.G.: *Can. J. Chem.* **51**, 1432 (1973)
4. Olah, G.A., Kuhn, S.J., Tolgyesi, W.S., Baker, E.B.: *J. Am. Chem. Soc.* **84**, 2733 (1962)
5. Olah, G.A., Tolgyesi, W.S., Kuhn, S.J., Moffatt, M.E., Bastien, I.J., Baker, E.B.: *J. Am. Chem. Soc.* **85**, 1328 (1963)
6. Olah, G.A., Comisarow, M.B.: *J. Am. Chem. Soc.* **88**, 4442 (1966)
7. Deno, N.C., Pittman, C.U., Wisotsky, M.J.: *J. Am. Chem. Soc.* **86**, 4370 (1964)
8. Veillard, A., IBMOL: Computation of wave functions for molecules of general geometry, version 4. IBM Research Laboratory, San Jose, California
9. Klessinger, M.: *Theoret. Chim. Acta (Berl.)* **15**, 353 (1969)
10. Costain, C.C., Dowling, J.M.: *J. Chem. Phys.* **32**, 158 (1960)
11. Hopkinson, A.C., Holbrook, N.K., Yates, K., Csizmadia, I.G.: *J. Chem. Phys.* **49**, 3596 (1968)
12. Hopkinson, A.C., Yates, K., Csizmadia, I.G.: *J. Chem. Phys.* **52**, 1784 (1970)
13. Jansen, H.B., Ros, P.: *Chem. Phys. Letters* **3**, 140 (1969)

14. Herzberg, G., Rao, K. N.: J. Chem. Phys. **17**, 1099 (1949)
15. Robb, M. A., Csizmadia, I. G.: Theoret. Chim. Acta (Berl.) **10**, 269 (1968)
16. Bell, R. P.: The proton in chemistry, p. 24. Ithaca: Cornell University Press 1959
17. Hammond, G. S.: J. Am. Chem. Soc. **77**, 334 (1955)
18. Duffy, J. A., Leisten, J. A.: J. Chem. Soc. **1960**, 853
19. Olah, G. A., Dunne, K., Mo, Y. K., Szilagyi, P.: J. Am. Chem. Soc. **94**, 4200 (1972)
20. Harrison, A. G., Ivko, A., van Raalte, D.: Can. J. Chem. **44**, 1625 (1966)
21. Tel, L. M., Wolfe, S., Csizmadia, I. G.: J. Chem. Phys. (in press)

Prof. Dr. A. C. Hopkinson  
Department of Chemistry  
York University  
Downsview, Ontario  
Canada



## *Relationes* An Application of RPA Theory to Conjugated Systems in the Excited States

### II. Linear Polyenes

Kazuo Kitaura and Kichisuke Nishimoto

Department of Chemistry, Osaka City University, Sumiyoshi-Ku, Osaka, Japan

Received January 15, 1973

In order to obtain a relationship between the molecular dimension and the correlation effect, RPA method has been applied to the calculation of electronic transition energies of linear polyenes. It has been found that the effect of electron correlation on the excitation energy decreases with increasing the size of molecule. The calculated oscillator strengths are remarkably improved by RPA calculation.

*Key words:* RPA calculation - Conjugated systems - Polyenes - Transition energies.

The calculation of the electronic transition energy is usually carried out by considering the configuration interaction (CI) among the excited configurations. It can be assumed that the configuration of very high energy would make no significant contribution to the lower excited states. Therefore, the CI procedure is usually made among only a limited number of the lower singly excited configurations. However, in some molecules, certain lower excited states are largely affected by doubly excited configurations [1]. For large molecules, these effects are not so clear, since the number of doubly excited configurations becomes very large.

The RPA method provides a tool of studying correlation effect on the lower excited states avoiding troublesome with large CI calculation [2, 3]. The RPA method does not handle doubly excited configurations explicitly, but it takes partly into account their effects to the lower excited states as de-excitation processes of doubly excited configurations included in the ground state. Previously, Donath [4] showed that the third excited state of benzene was strongly affected by the doubly excited configurations. As we have shown in the previous paper [5], the RPA calculation gave the same result. It is interesting to know the dependence of molecular size on the correlation effect to the excitation energy, because the electron repulsion integrals over molecular orbitals (MO's) are reduced as the dimensions of MO's increase. For this purpose, the energies of  $\pi - \pi^*$  transitions of linear polyenes are calculated by RPA method in this paper. The procedure of calculation and the parametrization are the same with those in our previous paper [5]. In the calculation, all of the C—C bond lengths and the bond angles were set to 1.4 Å and 120°, respectively. The two center core integral ( $\beta_{cc}$ ) was taken to be equal to -2.2 eV for all molecules. The electron repulsion integrals were calculated using Slater type AO with appropriate orbital exponent,  $\zeta$ .

As we have reported in the previous paper [5], for the RPA calculation, the orbital exponent calculated by Slater rule gave quite good results. Therefore, the same parametrization ( $\zeta_c = 1.625$ ) is used in this paper.

The results obtained by the method of configuration interaction among all possible singly excited configurations (SECI) and the corresponding RPA calculation are shown in Tables 1, 2 and Fig. 1, together with the experimental values. In these tables, only the results for the lower excited states are given.

In order to examine the influence of molecular geometry on the correlation effect, the singlet energy levels of *cis*- and *trans*-butadienes are calculated. The results are shown in Table 1. As seen from the table, the correlation effect in the *cis*-form seems to be rather larger than that in the *trans*-form. But, the difference is quite small. It is interesting to note that the first and third energy levels of *trans*-butadiene vary to a considerable extent by the RPA method, compared with those obtained by SECI calculation. For a comparison, another value of  $\zeta_c$ , which is determined as reproduce the value of Pariser-Parr approximation [6] for one center electron repulsion integral, i.e., 1.045 for carbon, is examined. The results calculated by this parameter are given in the same table. It should be noted that in this case the second transition of butadiene is rather largely varied by the mixing of de-excitation processes. Allinger [7] has reported that this transition was remarkably affected by the doubly excited configurations, when he calculated the transition energies using semi-empirical electron repulsion integrals. Namely the effect of doubly excited configurations (or de-excitation processes in the case of RPA method) depends largely on the electron repulsion integrals used, which we have already pointed out in the previous paper [5]. Semi-empirical electron repulsion integrals might already include partly a correlation effect. Therefore, when we use semi-empirical parameters, then the CI calculation including the doubly excited configurations may overestimate the correlation effect.

Table 1. Excitation energies ( $\Delta E$  in eV) and oscillator strengths (in paranthesis) of *cis*- and *trans*-butadienes

Case 1 <sup>a</sup>		Case 2 <sup>b</sup>		Obs	Ref.
SECI	RPA	SECI	RPA		
Cis-Butadiene					
5.99 (0.503)	5.04 (0.288)	5.07 (0.463)	4.96 (0.366)	4.95 <sup>c</sup>	[8]
7.26 (0.000)	7.22 (0.000)	7.91 (0.000)	7.89 (0.000)		
9.71 (1.114)	8.29 (0.534)	8.41 (0.933)	8.07 (0.669)		
9.78 (0.006)	9.62 (0.015)	10.41 (0.048)	10.38 (0.036)		
Trans-Butadiene					
6.44 (1.242)	5.71 (0.704)	5.53 (1.210)	5.36 (0.915)	5.71	[9]
7.50 (0.000)	7.48 (0.000)	7.93 (0.000)	7.68 (0.000)		
9.21 (0.000)	7.98 (0.000)	8.10 (0.000)	8.09 (0.000)		
10.06 (0.264)	9.86 (0.130)	10.64 (0.164)	10.62 (0.153)		

<sup>a</sup>  $\zeta_c = 1.625$  and  $\beta_{cc} = -2.2$  eV.

<sup>b</sup>  $\zeta_c = 1.045$  and  $\beta_{cc} = -2.8$  eV.

<sup>c</sup> The value of cyclohexadiene.

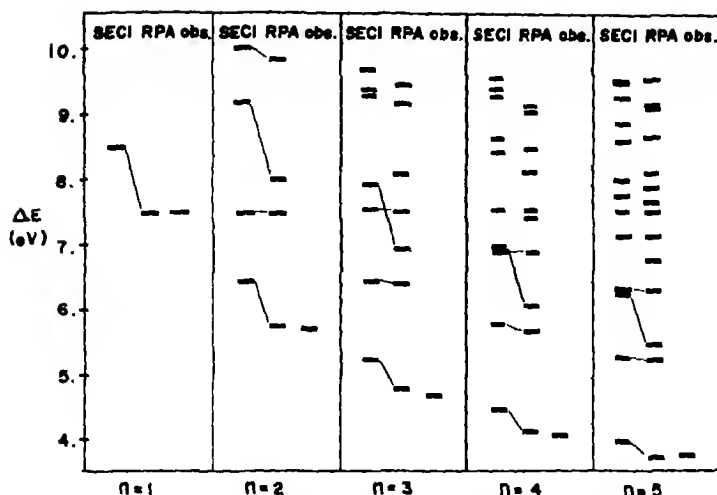


Fig. 1. Correlation diagram of excitation energies of linear polyenes

Table 2 and Fig. 1 show the lower singlet excitation energies and the oscillator strengths of linear polyenes calculated by the RPA and the SECI methods. The lowest excitation energies of small molecules are considerably lowered by the RPA treatment. However, as the size of molecule becomes larger, this effect becomes of little importance. This result shows that the correlation energy of the lowest excited state of large molecule cancels with that of the ground state within the RPA method. The reason for this may be due to a fact that an increase of the content of ionic structure (in terms of the terminology in valence bond theory) by an excitation becomes smaller as the dimension of MO becomes larger. It is evident that the electron correlation brings always to decrease such ionic structures.

As seen from Table 2, the calculated oscillator strengths are remarkably improved by RPA treatment. For example,  $f$  of *trans*-butadiene decreases from 1.242 to 0.704 which is very close to the experimental value ( $f=0.53$ ). It might be meaningful to estimate effective interactions between  $\pi$ -electrons in the excited states. The transition energy of ethylene can be expressed as follows in ZDO approximation;

$$\Delta E = -2\beta + (1/2)(\gamma_{11} - \gamma_{12}).$$

where  $\gamma_{\mu\nu}$  means the electron repulsion integral associated with  $2p\pi$  AO's,  $\phi_\mu$  and  $\phi_\nu$ . When we assume that the transition energy of ethylene calculated by the RPA method can be given by the same expression and we define the corresponding  $\gamma_{\mu\nu}$ 's as "effective" electron repulsion integrals,  $\gamma_{\mu\nu}^{\text{eff}}$ 's, then we obtain

$$\gamma_{11}^{\text{eff}} - \gamma_{12}^{\text{eff}} = 6.117 \text{ eV}.$$

In benzene, this quantity was to be 6.040 eV. It should be noted that this value is almost the same as the value calculated by our semi-empirical formula [12].

Table 2. Energies ( $\Delta E$  in eV) and oscillator strengths (in paranthesis) of the lower excited states of linear polyenes;  $H-(CH=CH)_n-H$ 

SECI	RPA	Obs.	Ref.
$n = 1$			
8.52 (0.730)	7.46 (0.377)	7.47	[10]
$n = 2$			
6.44 (1.242)	5.71 (0.704)	5.71	[9]
7.50 (0.000)	7.48 (0.000)		
9.21 (0.000)	7.98 (0.000)		
10.06 (0.264)	9.86 (0.130)		
$n = 3$			
5.27 (1.706)	4.78 (1.009)	4.71	[11]
6.47 (0.000)	6.43 (0.000)		
7.55 (0.000)	6.96 (0.000)		
7.95 (0.000)	7.54 (0.000)		
$n = 4$			
4.52 (2.124)	4.16 (1.312)	4.15	[11]
5.77 (0.000)	5.72 (0.000)		
6.93 (0.000)	6.14 (0.000)		
7.01 (0.000)	6.90 (0.000)		
$n = 5$			
4.02 (2.497)	3.74 (1.613)	3.80	[11]
5.31 (0.000)	5.24 (0.000)		
6.27 (0.000)	5.50 (0.000)		
6.36 (0.000)	6.33 (0.000)		

### References

1. Hirota, F., Nagakura, S.: *Bull. Chem. Soc. Japan* **43**, 1010 (1970)
2. Dunning, H. T., McKoy, V.: *J. Chem. Phys.* **47**, 1735 (1967)
3. McLachlan, A. D., Ball, M. A.: *Rev. Mod. Phys.* **36**, 844 (1964)
4. Donath, W. E.: *J. Chem. Phys.* **40**, 77 (1964)
5. Kitaura, K., Nishimoto, K.: *Theor. Chim. Acta (Berl.)* (in press)
6. Pariser, R., Parr, R. G.: *J. Chem. Phys.* **21**, 767 (1953)
7. Allinger, N. L., Tai, J. C.: *J. Amer. Chem. Soc.* **87**, 2081 (1965)
8. Schuler, H., Lutz, E., Arnold, G.: *Spectrochim. Acta* **17**, 1043 (1961)
9. Kierstead, R. W., Linstead, R. P., Weeden, B. C. L.: *J. Chem. Soc.* 1803 (1953)
10. Wilkinson, P. G., Johnston, H. L.: *J. Chem. Phys.* **18**, 190 (1950)
11. Bohlmann, F., Mannhardt, H.: *Chem. Ber.* **89**, 1307 (1956)
12. Nishimoto, K., Mataga, N.: *Z. Physik. Chem. (Frankfurt)* **12**, 335 (1957)

Prof. Dr. K. Nishimoto  
 Department of Chemistry,  
 Osaka City University  
 Sumiyoshi-Ku, Osaka  
 Japan



## CNDO/2 and INDO Calculations of a Reaction Pathway for the Sigmatropic [1, 5] H-Shift in Cyclopentadiene

J. R. de Dobbelaere\*, J. W. de Haan\*\*, H. M. Buck\*, and G. J. Visser\*\*\*

Laboratories of Organic Chemistry\* and of Instrumental Analysis\*\* and Computing Centre\*\*\*,  
University of Technology, Eindhoven, The Netherlands

Received October 16, 1972/April 6, 1973

CNDO/2 and INDO calculations have been carried out in order to construct a suitable model for the activated complex during the reaction. In this reaction model the migrating hydrogen atom moves along an edge of the cyclopentadiene ring. An analysis of this situation suggests a partial electron transfer from the migrating hydrogen to the nascent cyclopentadienyl system. This charge transfer is discussed in terms of aromaticity. The calculated activation enthalpies are 10 kcal/mole (CNDO/2) and 17 kcal/mole (INDO), whereas the experimental value is ca. 24 kcal/mole [1].

**Key words:** Thermal sigmatropic hydrogen shifts — Transition state geometry — Aromaticity of odd-membered cycloradicals

### Introduction

Thermal [1, 5] sigmatropic rearrangements have been reported in many linear and cyclic conjugated polyenes [2, 3, 4, 5]. The kinetic activation parameters of [1, 5] H-shifts show large mutual differences,  $\Delta H^*$  varying between ca. 20 and 40 kcal/mole.

Starting an attempt to correlate these energy differences with transition state geometries, we performed CNDO/2 and INDO calculations according to Pople [6, 7].

In a similar way modified INDO calculations have been described earlier for the Cope rearrangement [8] and the butadiene  $\rightleftharpoons$  cyclobutene isomerization [9]. Very recently the activation energy for the suprafacial and antarafacial [1, 5] H-shift in *cis*-piperylene has been calculated using the MINDO/CI method [10].

### Results

In the transition state of the [1, 5] sigmatropic H-shift in cyclopentadiene we assumed a three-center bond between the migrating hydrogen atom and two carbon atoms. This presupposition was justified by the results of Shchembelov and Ustynyuk who considered also the four- and six-center bonds [11]. The ground state geometry was constructed using microwave spectroscopic data [12]. Concerning the bond lengths and angles which were varied in our minimization procedure, we used the values obtained by Shchembelov and Ustynyuk [11], see Fig. 1.

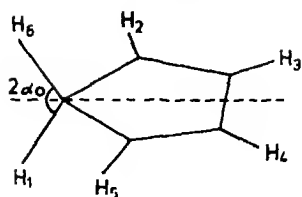


Fig. 1

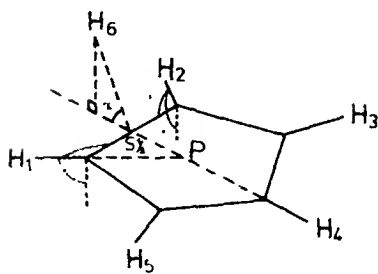


Fig. 2

Assuming a suprafacial sigmatropic shift of  $H_6$  from  $C_1$  to  $C_2$ , the symmetric transition state will be as depicted in Fig. 2. In this study the reaction coordinate is defined by the angle  $(S_R PC_1)$ .

R	$(S_R PC_1)$	
1.8	10	$S_R \cdot P$ and $H_6$ are situated in a plane perpendicular to the cyclopentadiene ring, while $S_R$ moves along $C_1 - C_2$ .
1.4	20	
3.8	30	
1.2	40	

Energy minimization was carried out using the SIMPLEX method [13], with respect to eight geometrical parameters:

- angle  $\alpha$
- distance  $S_R - H_6$
- angles  $\beta$  and  $\gamma$  (corresponding with out-of-plane movements of  $H_1$  and  $H_2$ )
- angles  $\delta$  and  $\epsilon$  (in the cyclopentadiene plane)
- distance  $C_1 - C_2$  (two parameters).

Other dimensions have not been optimized by CNDO/2 or INDO methods. They were derived from the ground state dimensions by applying simple geometric relations. See Table 1.

Table 1

	R = 0	R = 1.8	R = 1.4	R = 3.8	R = 1/2	R = 1
$C_3 - C_4$ (Å)	1.460	1.450	1.441	1.411	1.402	1.344
$C_4 - C_5$ (Å)	1.344	1.366	1.370	1.383	1.399	1.460
$C_3 C_4 C_5$ (°)	109.3	109.1	108.9	108.4	108.6	109.3

Table 2. CNDO/2 results

R	C <sub>1</sub> -C <sub>2</sub>	S <sub>R</sub> -H <sub>6</sub> (Å)	$\alpha$ (°)	$\beta$ (°)	$\gamma$ (°)	$\delta$ (°)	$\varepsilon$ (°)	C <sub>1</sub> C <sub>2</sub> C <sub>3</sub> (°)	$\Delta E$ (kcal/mole)	$\rho H_6$
0	1.470	1.090	53.0	-53.0	0.0	127.3	124.4	108.3	0.0	0.97
1/8	1.465	1.138	63.3	-36.2	-1.9	127.7	122.7	109.8	6.0	0.95
1/4	1.460	1.119	71.0	-25.9	-2.2	129.4	123.5	109.6	10.0	0.90
3/8	1.460	1.075	76.1	-13.4	-3.2	127.6	123.3	108.4	10.4	0.85
1/2	1.470	1.050	76.6	-6.3	-6.9	126.7	125.5	107.6	10.0	0.83

Table 3. INDO results

R	C <sub>1</sub> -C <sub>2</sub> (Å)	S <sub>R</sub> -H <sub>6</sub> (Å)	$\alpha$ (°)	$\beta$ (°)	$\gamma$ (°)	$\delta$ (°)	$\varepsilon$ (°)	C <sub>1</sub> C <sub>2</sub> C <sub>3</sub> (°)	$\Delta E$ (kcal/mole)	$\rho H_6$
0	1.470	1.090	53.0	-53.0	0.0	127.3	124.4	108.3	0.0	1.00
1/8	1.470	1.146	62.9	-36.8	1.8	127.6	122.7	109.7	6.7	0.98
1/4	1.463	1.130	69.6	-26.6	-1.4	128.6	123.6	109.5	12.9	0.93
3/8	1.468	1.079	74.4	-13.8	-2.7	126.4	124.5	108.2	16.5	0.88
1/2	1.471	1.058	76.0	-8.8	-5.8	125.9	125.3	107.0	17.2	0.86

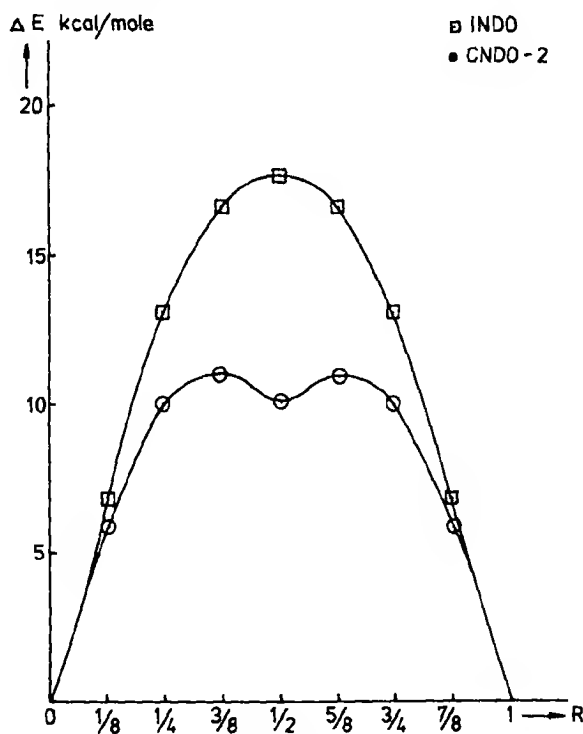


Fig. 3

In fact, variation of these dimensions is bound to have some influence on the total energy. However, a very similar influence is to be expected in the total energy of the ground state molecule, thus the energy difference is not affected.

The resulting energies, geometries and electron densities on the migrating hydrogen atom are shown in Tables 2 and 3 and Fig. 3.

### Discussion

The results of the CNDO/2 and INDO calculations for the reaction pathway indicate that the electron density on the migrating hydrogen varies between 1.0 and 0.8. This is immediately reminiscent of one of Woodward and Hoffmann's early remarks: "In the  $[1, j]$  sigmatropic migration of hydrogen within an all-*cis*-polyene framework,  $R_2C^1=CH-(CH=CH)_k-C^jHR_2$ , one may envisage the transition state as made up by the combination of the orbital of a hydrogen atom with those of a radical containing  $2k + 3\pi$ -electrons".

Generally, at each point of the reaction coordinate a charge transfer process takes place which can be described as a "hybrid" of two configurations. The intermediate position ( $R = 1/2$ ) can be indicated as follows:



In this situation all carbon atoms are  $sp^2$ -hybridized, thus creating optimal conditions for charge delocalization. In the CNDO/2 approximation this situation is calculated to be an intermediate. This intermediate is probably inherent to the CNDO/2 method, it does not appear in the INDO-calculations. The weights for the "no-bond" configuration and the "dative" configuration are approximately in the ratio 6 : 1. The relatively small contribution of the latter "aromatic" configuration is due to the high ionization potential of hydrogen and its small penetration magnitude. If one uses aromaticity as a criterion for predicting the course of the reaction, the question whether radicals containing  $2n + 3$  conjugated carbon atoms and  $2n + 3\pi$ -electrons are aromatic, should be solved first. Dewar suggests that Hückel  $4n + 3$ -radicals are probably aromatic although there might be some doubt about the cyclopropenyl radical [14]<sup>1</sup>.

Simple PPP calculations indeed predict aromaticity for the cyclopropenyl and cyclopentadienyl radicals, compared with the open structure radicals. The aromaticity of the cycloheptatrienyl radical has been concluded after resonance energy determining experiments [15].

We might conclude that in our picture a suprafacial hydrogen shift is only thermally allowed when both the "dative" and the "no-bond" configurations are aromatic.

The reason why " $4n + 2$ " aromaticity may be used as a criterion for an allowed sigmatropic process in cyclic systems, seems to be that cycloradicals containing  $2n + 3$  carbon atoms and  $2n + 3\pi$ -electrons are aromatic.

The reaction pathway borne out by our calculations may be visualized as follows: between  $R = 0$  and  $R = 1/2$  the atom  $H_1$  shifts gradually towards the plane of the cyclopentadiene ring, while  $H_2$  will stay in or close to that plane ( $C_2$   $sp^2$ -hybridized). This leads to a symmetrical state with both  $H_1$  and  $H_2$  close to (within 0.1 Å) the plane of the ring when  $R = 1/2$  ( $C_1$  and  $C_2$   $sp^2$ -hybridized). The angle  $\alpha$  increases gradually from 53 to 77°.

<sup>1</sup> In this respect it may be of interest to note that (+)-1-bromo-2-methylbutane reacts with  $Br^-$  and  $DBr$  under formation of (+)-1-bromo-2-methyl-2-D-butane with retention of configuration [16].

This discussion is substantiated by preliminary calculations on the 1,3-cyclohexadiene system. We find analogous parameters with ringbending as an extra feature. The electron density on the migrating hydrogen ranges between 0.8 and 0.9 in the transition state. Mechanistic implications of the shifts in the cyclohexadiene ring are apparently very similar to those described for the cyclopentadiene system. A detailed description and kinetic parameters of the 1,3-cyclohexadiene system will be published in due course.

### References

1. Roth, W.R.: *Tetrahedron Letters* **1964**, 1009
2. Mironov, V.A., Sobolev, E.B., Elizarova, A.N.: *Tetrahedron* **19**, 1939 (1963)
3. Roth, W.R., König, J.: *Liebigs Ann. Chem.* **699**, 24 (1966)
4. ter Borg, A.P., Razenberg, E., Kloosterziel, H.: *Rec. Trav. Chim.* **84**, 1230 (1965)
5. de Haan, J.W., Kloosterziel, H.: *Rec. Trav. Chim.* **87**, 298 (1968)
6. Pople, J.A., Santry, D.P., Segal, G.A.: *J. Chem. Phys. Suppl.* **43**, S 129 (1965)  
Segal, G.A.: *J. Chem. Phys. Suppl.* **43**, S 136 (1965)
7. Pople, J.A., Beveridge, D.L., Dobosch, P.A.: *J. Chem. Phys.* **47**, 2026 (1967)
8. Dewar, M.J.S., Lo, D.H.: *J. Am. Chem. Soc.* **93**, 7201 (1971)
9. McIver Jr., J.W., Komornicki, A.: *J. Am. Chem. Soc.* **94**, 2625 (1972)
10. Bingham, R.C., Dewar, M.J.S.: *J. Am. Chem. Soc.* **94**, 9107 (1972)
11. Shchembelov, G.A., Ustynyuk, Yu.: *Theoret. Chim. Acta (Berl.)* **24**, 389 (1972)
12. Sharpe, L.H., Laurie, V.H.: *J. Chem. Phys.* **43**, 2760 (1965)
13. Nelder, J.A., Mead, R.: *Computer J.* **7**, 308 (1964)
14. Dewar, M.J.S., Karschner, S.: *J. Am. Chem. Soc.* **93**, 4290 (1971)
15. Vincow, G., Dauben, H.J., Hunter, F.R., Volland, W.V.: *J. Am. Chem. Soc.* **91**, 2823 (1969)
16. Shea, K.J., Skell, P.S.: *J. Am. Chem. Soc.* **95**, 283 (1973)

J. R. de Dobbelaere  
Laboratory of Organic Chemistry  
University of Technology  
Eindhoven, The Netherlands



## Oscillations in the Thomas-Fermi-Dirac Electron Density

Jerry Goodisman

Department of Chemistry, Syracuse University, Syracuse, New York

Received May 17, 1973

It is shown that there can be no local maxima (except at nuclei) in the electron density predicted by models, such as the Thomas-Fermi-Dirac theory, which make the electron density a monotonic function of the electrostatic potential.

**Key words:** Density oscillations – Thomas-Fermi-Dirac model.

It is well known [1] that the Thomas-Fermi and Thomas-Fermi-Dirac models predict, for atoms, an electron density lacking the oscillations associated with shell structure. We have been interested in assessing the utility of these models for calculation of molecular electron densities. In this note, we show that, for a molecule, these models cannot give rise to local maxima in the electron density, or to points in space at which the electron density is constant along one or two perpendicular spatial directions and a maximum in those remaining, except at a nucleus (where the Thomas-Fermi-Dirac electron density is infinite). Hohenberg and Kohn [2] have discussed the impossibility of density fluctuations by considering properties of the polarizability in momentum space.

Here, we prove the impossibility of local maxima from the condition that the electron density  $\rho$  be a monotonic function of the electrostatic potential  $\Phi$ . In the Thomas-Fermi-Dirac theory

$$\rho = f(\Phi) = \left[ \frac{2\kappa_a}{5\kappa_k} + \left( \frac{-3e\Phi}{5\kappa_k} + \frac{16\kappa_a^2}{100\kappa_k^2} \right)^{1/3} \right]^3$$

where  $\kappa_a$  and  $\kappa_k$  are constants. It is easily seen that  $f(\Phi)$  is monotonic.

Suppose that, at some point where  $\rho$  is finite, one has

$$\partial\rho/\partial x = \partial\rho/\partial y = \partial\rho/\partial z = 0 \quad (a)$$

and

$$\partial^2\rho/\partial x^2 \leq 0; \quad \partial^2\rho/\partial y^2 \leq 0; \quad \partial^2\rho/\partial z^2 \leq 0. \quad (b)$$

Since  $\partial\rho/\partial x = (df/d\Phi)(\partial\Phi/\partial x)$ , we have

$$\partial\Phi/\partial x = \partial\Phi/\partial y = \partial\Phi/\partial z = 0.$$

Since

$$\frac{\partial^2\rho}{\partial x^2} = \frac{d^2f}{d\Phi^2} \left( \frac{\partial\Phi}{\partial x} \right)^2 + \frac{df}{d\Phi} \left( \frac{\partial^2\Phi}{\partial x^2} \right)$$

and the first term vanishes, the conditions (b) imply

$$\frac{df}{d\Phi} \left( \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} \right) \leq 0.$$

But  $df/d\Phi$  is positive, and  $\nabla^2 \Phi$  is positive since the potential satisfies the Poisson equation.

Therefore, (a) and (b) cannot hold. Thus, local maxima in  $\varrho$  are impossible (except at the nuclei), and there can be no density fluctuations.

### References

1. Gonbas, P., *Die statistische Theorie des Atoms und ihre Anwendungen*. Vienna: Springer 1949.
2. Hohenberg, P., Kohn, W.: *Phys. Rev.* **136**, B 864 (1964)

Prof. Dr. J. Goodisman  
Department of Chemistry  
Syracuse University  
Syracuse 13210, New York  
USA



## Commentationes

# Three-Dimensional Bond-Charge Models for Potential Curves of Diatomic Molecules\*

George A. Henderson\*\* and Robert G. Parr

Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218

Received May 22, 1973

Several simple three-dimensional Fermi-gas models for potential energy curves of diatomic molecules are suggested. Bond-charge parameters close to those predicted by the earlier point bond-charge model of Borkman, Simons and Parr [J. Chem. Phys. 49, 1055 (1968); 50, 58 (1969)] are obtained for models assuming uniform spherical or elliptical electron distributions in the bond region.

**Key words:** Potential curves   Bond-charge - Fermi-gas

## 1. Introduction

The point of departure for this work is the simple bond-charge model for potential curves of diatomic molecules described by Borkman, Simons and Parr [1]. They considered a point charge of magnitude  $-qe$  restricted to motion in one dimension over a length  $\nu R$ , where  $R$  is the internuclear separation. The kinetic energy of such a charge varies as  $1/R^2$  and the electrostatic interaction with two point-charge atoms of charge  $Ze = \frac{1}{2}qe$  gives a potential energy varying as  $1/R$ . This startlingly simple model gives rise to a total electronic energy of the form:

$$W(R) = T(R) + V(R) = W_0 + W_1/R + W_2/R^2, \quad (1)$$

which reasonably well represents experimental diatomic potential curves near equilibrium. Further, the values of  $q$  which result from fitting the model to experiment can be interpreted as bond orders.

Here, we investigate three other simple models, two of which also yield the form Eq. (1). The new models all involve three-dimensional charge distributions. They also employ the assumption that the electronic kinetic energy depends upon the electron density in the same way as in a statistical or Thomas-Fermi theory. That is, a kinetic energy density proportional to the  $5/3$  power of the density is assumed.

We consider that at every point in coordinate space, momentum states are fully occupied up to the Fermi momentum  $P_F = \hbar k_F$ , as in a degenerate Fermi

\* Aided by a research grant to The Johns Hopkins University from the National Science Foundation.

\*\* Present address: Department of Physics, Southern Illinois University, Edwardsville, Illinois 62025.

gas [2]. This gives the fundamental density-momentum relation

$$\begin{aligned} \varrho(r) &= \sum_s \int d\mathbf{k} \varphi_s^*(r, s) \varphi_s(r, s) \\ &= \frac{2}{(2\pi^3)} 4\pi \int_0^{k_F(r)} dk k^2 \\ \varrho(r) &= \frac{1}{3\pi^2} k_F^3(r). \end{aligned} \quad (2)$$

The average (kinetic) energy at a point in space is given by

$$T(r) = \frac{\int dE E N(E)}{\int dE N(E)}, \quad (3)$$

where  $N(E)$  is the energy distribution function for a non-interacting Fermi gas confined to a volume  $V$  and equals

$$N(E) dE = \frac{m^{3/2} V}{2^{1/2} \pi^2 \hbar^3} E^{1/2} dE, \quad (4)$$

where  $m$  is the particle mass. Carrying out the integrations and expressing the result in terms of the Fermi momentum, one obtains

$$T(r) = \frac{3}{5} \frac{k_F^2(r)}{2m}. \quad (5)$$

Thus the total kinetic energy is given by

$$\begin{aligned} T &= \int d\mathbf{r} \varrho(r) T(r) \\ &= \frac{3\hbar^2}{10m} \left( \frac{3}{8\pi} \right)^{2/3} \int d\mathbf{r} \varrho^{5/3}(r), \end{aligned} \quad (6)$$

where Eq. (2) has been used. This relation is assumed for all the models discussed here.

## 2. Spherical Charge Density Model

The first model assumes that the bond charge  $-eq$  in a homonuclear diatomic molecule is uniformly distributed in a sphere of radius  $\alpha R/2$  located at the center of the bond, while point charges  $Ze = \frac{1}{2}qe$  are located at each of the atoms. Thus for the electron density we have

$$\varrho(r) = \varrho_0 \Theta(\alpha R/2 - r), \quad (7)$$

where  $\varrho_0$  is a constant,  $r$  is the radial variable in a spherical coordinate system located at the bond center and

$$\Theta(x) = \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0. \end{cases} \quad (8)$$

The constant  $\varrho_0$  is determined by normalization:

$$q = \int d\mathbf{r} \varrho(r) = \frac{\varrho_0 \pi \alpha^3 R^3}{6}. \quad (9)$$

Computing the kinetic energy via Eq. (6), we obtain

$$T = \frac{W_2}{R^2} = \frac{h^2}{2m} \left( \frac{3}{2\pi} \right)^{4/3} \frac{q^{5/3}}{\alpha^2 R^2}. \quad (10)$$

For the electronic potential energy we obtain

$$V = \frac{W_1}{R} = \begin{cases} \frac{e^2}{R} (Z^2 - 4Zq) & \text{for } \alpha < 1, \\ \frac{e^2}{R} \left( Z^2 - \frac{2Zq}{\alpha^3} [3\alpha^2 - 1] \right) & \text{for } \alpha > 1. \end{cases} \quad (11)$$

The case  $\alpha < 1$  is equivalent from the point of view of potential energy to the point-charge Borkman-Simons-Parr model. When  $\alpha$  exceeds 1, the bond charge density extends beyond the positions of the two atoms and different values of  $q$  result. The two parameters  $q$  and  $\alpha$  are determined from  $R_e$ , the experimental equilibrium distance and  $k_e$ , the quadratic force constant:

$$\left. \frac{dW}{dR} \right|_{R_e} = 0; \quad \left. \frac{d^2 W}{dR^2} \right|_{R_e} = k_e. \quad (12)$$

Values of  $q$  and  $\alpha$  for the ground states of 17 homonuclear diatomic molecules are presented in Table 1. Figure 1 permits a comparison of these values of  $q$  to those obtained from the Borkman-Simons-Parr one-dimensional model and to the simple bond order. It will be seen that the bond charges from the spherical Fermi-gas model are in qualitative agreement with the one-dimensional values but tend to be a bit higher throughout, especially in the neighborhood of  $N_2$ . Note also that all values of  $\alpha$  exceed 1.

Table 1. Bond charge and size parameters for spherical Fermi-gas model<sup>a</sup>

Molecule	$q$	$\alpha$
H <sub>2</sub>	1.78	5.24
Li <sub>2</sub>	1.15	1.33
B <sub>2</sub>	2.16	1.70
C <sub>2</sub>	2.79	1.88
N <sub>2</sub>	3.49	2.08
O <sub>2</sub>	2.84	2.03
F <sub>2</sub>	2.17	1.87
Na <sub>2</sub>	1.14	1.21
Si <sub>2</sub>	2.52	1.24
P <sub>2</sub>	3.21	1.33
S <sub>2</sub>	3.04	1.35
Cl <sub>2</sub>	2.68	1.34
K <sub>2</sub>	1.21	1.03
Se <sub>2</sub>	3.06	1.23
Br <sub>2</sub>	2.76	1.21
Te <sub>2</sub>	3.20	1.06
I <sub>2</sub>	2.85	1.09

<sup>a</sup> Model of Eqs. (10) and (11) of text. Bond-charge parameter  $q$  and sphere-size parameter  $\alpha$  determined from Eq. (12) of text.

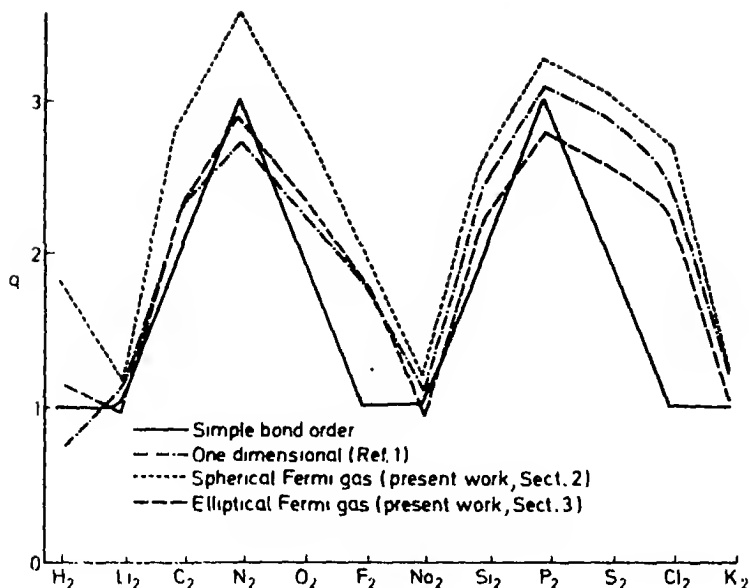


Fig. 1

### 3. Elliptical Charge Density Model

A potential energy  $W(R)$  of the form Eq. (1) also results when one assumes a uniform charge distribution within an ellipsoidal volume. This more nearly represents the shape of a molecular electronic charge distribution. We employ elliptical [prolate spheroidal] coordinates  $\xi = \frac{r_A + r_B}{R}$ ,  $\eta = \frac{r_A - r_B}{R}$ , and  $\varphi$ , where  $r_A$  and  $r_B$  are distances from nuclei A and B, respectively, to an arbitrary point in space and  $\varphi$  is an angle of rotation about the bond axis. Since curves of constant  $\xi$  are ellipses with foci at atoms A and B, we can employ the step function to cut off the density at a value  $\xi = \beta$ . Thus we have

$$\varrho(r) = \varrho_0 \Theta(\beta - \xi), \quad (13)$$

and the charge normalization condition takes the form

$$\begin{aligned} q &= \varrho_0 \int_0^{2\pi} d\varphi \int_{-1}^1 d\eta \int_1^\beta d\xi (\xi^2 - \eta^2) \left(\frac{R}{2}\right)^3 \\ &= \frac{\varrho_0 \pi R^3}{6} (\beta^3 - \beta), \end{aligned} \quad (14)$$

which determines  $\varrho_0$ . Proceeding as before, we obtain the electronic kinetic energy expression *via* Eq. (4), yielding

$$T = \frac{W_2}{R^2} = \frac{3h^2}{10m} \left[ \frac{9}{4\pi^2(\beta^3 - \beta)} \right]^{2/3} \frac{q^{5/3}}{R^2}. \quad (15)$$

Table 2. Bond charge and size parameters for elliptical Fermi-gas model\*

Molecule	$q$	$\beta$
H <sub>2</sub>	1.13	2.9
Li <sub>2</sub>	0.99	1.26
B <sub>2</sub>	1.82	1.43
C <sub>2</sub>	2.31	1.51
N <sub>2</sub>	2.85	1.61
O <sub>2</sub>	2.33	1.58
F <sub>2</sub>	1.80	1.50
Na <sub>2</sub>	0.99	1.21
Si <sub>2</sub>	2.19	1.23
P <sub>2</sub>	2.77	1.26
S <sub>2</sub>	2.62	1.27
Cl <sub>2</sub>	2.31	1.27
K <sub>2</sub>	1.04	1.14
Se <sub>2</sub>	2.65	1.22
Br <sub>2</sub>	2.39	1.21
Te <sub>2</sub>	2.76	1.16
I <sub>2</sub>	2.46	1.17

\* Model of Eqs. (15) and (16) of text. Bond-charge parameter  $q$  and ellipsoid-size parameter  $\beta$  determined from Eq. (12) of text.

The potential energy  $V_{A_e}$  due to the interaction of the negative charge distribution with the atomic charge at A is

$$\begin{aligned}
 V_{A_e} &= -Ze \int dr \frac{\rho(r)}{r_A} = -Ze^2 \rho_0 \int_0^{2\pi} d\varphi \int_{-1}^1 d\eta \int_1^\beta d\xi \left(\frac{R}{2}\right)^2 (\xi - \eta) \\
 &= \frac{-Ze^2 \rho_0 \pi R^2}{2} (\beta^2 - 1).
 \end{aligned}
 \tag{16}$$

Atom B experiences the same potential. When one adds the atomic repulsion term, the resulting expression for the potential energy in this model is

$$V = \frac{W_1}{R} = \frac{e^2}{R} \left( Z^2 - \frac{6Zq}{\beta} \right).
 \tag{17}$$

The values of  $q$  and  $\beta$  obtained by fitting the experimental  $R_e$  and  $k_e$  are displayed in Table 2. Note that the values of  $\beta$  are small enough that (except for H<sub>2</sub>) the charge is held rather close to the bond axis. Reference to Fig. 1 shows that the elliptical model gives values of  $q$  very close to those obtained by Borkman, Simons and Parr.

#### 4. Elliptical Charge Density Model with Ion Core

As an elaboration of the foregoing two-parameter models, one may consider a three-parameter model which accounts for the finite size of the ion cores by excluding a spherical volume about each atom center from an ellipsoidal charge density allotted to the electrons. Thus the bond charge distribution is constant in an ellipsoidal volume except for two spherical "bubbles" of radius  $\zeta$  within

which it vanishes and it also vanishes outside the ellipsoidal boundary defined by  $\xi = \beta$ . In terms of step functions, the density for this model is

$$\begin{aligned}\varrho(\xi, \eta) &= \varrho_0 \Theta(\beta - \xi) \Theta(r_A - \xi) \Theta(r_B - \xi) \\ &= \varrho_0 \Theta(\beta - \xi) \Theta(\tfrac{1}{2}R[\xi + \eta] - \zeta) \Theta(\tfrac{1}{2}R[\xi - \eta] - \zeta).\end{aligned}\quad (18)$$

The step functions give integration limits such that

$$\begin{aligned}1 &\leq \xi \leq \beta; \\ b - \xi &\leq \eta \leq \xi - b, \quad b \equiv 2\zeta/R.\end{aligned}\quad (19)$$

The bond-charge normalization condition is then

$$\begin{aligned}\int \varrho d\tau &= q = \varrho_0 2\pi \left(\frac{R}{2}\right)^3 \int_1^\beta d\xi \int_1^\xi d\eta (\xi^2 - \eta^2) \Theta(\beta - \xi) \Theta(r_A - \xi) \Theta(r_B - \xi) \\ &= \varrho_0 2\pi \left(\frac{R}{2}\right)^3 \left\{ \int_1^\beta d\xi \int_b^{\xi-b} d\eta (\xi^2 - \eta^2) + \int_{1+b}^\beta d\xi \int_{-1}^1 d\eta (\xi^2 - \eta^2) \right\}, \\ q &= \varrho_0 2\pi \left(\frac{R}{2}\right)^3 \left\{ \frac{2}{3} [\beta^3 - \beta] - \frac{4}{3} b^3 \right\}.\end{aligned}\quad (20)$$

Proceeding as before, the following expression for the electronic kinetic energy is obtained:

$$T = \frac{3h^2}{10m} \left( \frac{9}{4\pi^2} \right)^{2/3} \frac{q^{5/3}}{R^2} ([\beta^3 - \beta] - 2b^3)^{2/3}. \quad (21)$$

When the electronic-nuclear attraction potential is evaluated and the ion-core repulsion term is added, one obtains

$$V = \frac{e^2 Z^2}{R} - \frac{6Zq e^2}{R} \left[ \frac{\beta^2 - 1 - b^2 - 1/3 b^3}{\beta^3 - \beta - 2b^3} \right]. \quad (22)$$

The denominators in the  $T$  and  $V$  expressions can be expanded in binomial series to give

$$\begin{aligned}T &= \frac{3h^2}{10m} \left[ \frac{9}{4\pi^2} (\beta^3 - \beta) \right]^{2/3} \frac{q^{5/3}}{R^2} \left[ 1 + \frac{32}{3(\beta^3 - \beta)} \left( \frac{\zeta}{R} \right)^3 - \frac{640}{3(\beta^3 - \beta)^2} \left( \frac{\zeta}{R} \right)^6 + \dots \right], \\ V &= \frac{Z^2 e^2}{R} - \frac{6Zq e^2}{\beta R} \left[ 1 - \frac{4}{(\beta^2 - 1)} \left( \frac{\zeta}{R} \right)^2 + \frac{(6 - \beta)}{(\beta^3 - \beta)} \frac{8}{3} \left( \frac{\zeta}{R} \right)^3 - \dots \right].\end{aligned}\quad (23)$$

Thus this model includes higher powers of  $(1/R)$  than occur in the earlier models. In particular, the presence of a  $(1/R)^3$  term represents a more realistic description of the potential than does the form Eq. (1). One could now proceed to determine  $\beta$ ,  $q$ , and  $\zeta$  by fitting to the experimental equilibrium distance,  $R_e$ , and the quadratic and cubic force constants  $k_e$  and  $l_e$ . However, any physically intuitive interpretation of their significance would be partially vitiated by the fact that this model involves a partial contradiction to the virial theorem. If one postulates the general form

$$W(R) = \sum_{n=0}^{\infty} W_n R^{-n}, \quad (24)$$

the virial theorem predicts that

$$\begin{aligned} -T &= W + R \frac{dW}{dR} = W_0 - \frac{W_2}{R^2} - \frac{2W_3}{R^3} \dots - \frac{(n-1)W_n}{R^n} - \dots, \\ V &= 2W + R \frac{dW}{dR} = 2W_0 + \frac{W_1}{R} - \frac{W_3}{R^3} \dots - \frac{(n-2)W_n}{R^n} - \dots. \end{aligned} \quad (25)$$

Thus while the present three-parameter model gives a correctly behaving energy insofar as  $V$  has no  $1/R^2$  term and  $T$  has no  $1/R$  term, it falls short in that  $T$  has no  $1/R^3$  term.

### 5. Conclusions

A point of diminishing returns generally is reached when one complicates a simple model too much. Of the models we have here considered we prefer the two-parameter elliptical model of Section 3. Results with it demonstrate that the Borkman-Simons-Parr model can successfully be extended to three dimensions without significant change in the values of (or the interpretation of) the bond charge parameter.

In this work, in order to retain maximum simplicity in extending the point bond-charge model to three dimensions, a uniform density distribution has been assumed. Exploring a more realistic distribution, perhaps closer to a molecular Thomas-Fermi density, could be profitable. Further, the success of Simons [3] in predicting force constants for triatomic molecules using a point charge-point dipole model suggests that suitably simple three-dimensional Fermi-gas models for vibrating polyatomic molecules would be worth examining.

### References

1. Parr, R. G., Borkman, R. F.: *J. Chem. Phys.* **46**, 3683 (1967); **49**, 1055 (1968); Borkman, R. F., Parr, R. G.: *J. Chem. Phys.* **48**, 1116 (1968); Borkman, R. F., Simons, G., Parr, R. G.: *J. Chem. Phys.* **50**, 58 (1969). This model has recently been shown to be a first approximation to an exact theory of molecular vibrations near equilibrium. Cf. Simons, G., Parr, R. G.: *J. Chem. Phys.* **55**, 4197 (1971)
2. For example, Eisberg, R. M.: *Foundations of modern physics*. New York: John Wiley & Sons, Inc. 1962
3. Simons, G.: *J. Chem. Phys.* **56**, 4310 (1972)

Professor Robert G. Parr  
Department of Chemistry  
The Johns Hopkins University  
Baltimore, Maryland 21218, USA





## A Theory of Molecules in Molecules

### II. The Theory and Its Application to the Molecules Be-Be, Li<sub>2</sub>-Li<sub>2</sub>, and to the Internal Rotation in C<sub>2</sub>H<sub>6</sub>

W. von Niessen

Lehrstuhl für Theoretische Chemie, Technische Universität München, 8 München 2, Germany

Received February 20, 1973/May 23, 1973

A theory of molecules in molecules is presented, which permits the computation of the wave function of a molecule from the wave functions of fragment molecules by transferring some of the localized molecular orbitals of the fragments and recalculating the orbitals in the region of interaction. A projection operator is used to obtain orthogonality of the orbitals to be determined to the transferred and fixed orbitals. Additional approximations allow the reduction of the dimension of the matrices to be diagonalized and the neglect of a part of the basic integrals, which can lead to a considerable saving in the computation time. The justification of these approximations will be investigated for the case of the molecules Be-Be, Li<sub>2</sub>-Li<sub>2</sub>, and for the calculation of the rotational barrier in C<sub>2</sub>H<sub>6</sub>.

*Key words:* Localized orbitals, transferability of  $\sim$  - Be<sub>2</sub> - Li<sub>4</sub> - C<sub>2</sub>H<sub>6</sub> - Molecules in molecules

#### 1. Introduction

The Schrödinger equation is the fundamental equation of molecular quantum mechanics; it is an exact equation, if relativistic effects are neglected. The equation itself as well as its solutions describe nature in a very abstract way and the connection to the terms used to describe molecular structure as there are atoms, bonds, inner shells or lone pairs of electrons etc. can only indirectly be established. A similar statement can be made of the most frequently used approximation to the solution of the Schrödinger equation, the Hartree-Fock (HF) approximation. The HF equation and its solutions are abstract and lack the direct interpretability in the terms mentioned above. But it is possible to reformulate the Schrödinger equation and the HF equation so that their solutions can be interpreted more directly in familiar terms and thus appeal stronger to physical and chemical intuition. It is possible to make this reformulation in such a way that the solutions to the new equations are equivalent to the solutions of the original ones. Solving the new equations - which depend on the introduction of a model - does consequently not necessarily involve a loss in rigour. In most cases the reformulated equations are more difficult to solve than the original ones. From a mathematical point of view it is only in the cases where convergence difficulties arise that the new equations can be preferable, because the starting point is known. The main advantage of the reformulation is that physical and chemical intuition can be used to find solutions to the equations and to establish approximations to them. These approximations should be physically appealing and intelligible, be

justifiable in a precise manner, and the approximated equations should be easier to solve than the equations which formed the starting point. Quite a significant amount of work has been done in this direction. A number of approaches, which are of concern to the present work, will be shortly mentioned.

Pseudopotential theory [1–4] has been developed to remove orthogonality constraints in the calculation of wave functions for some subsystem of the electrons (e.g. the calculation of the valence electron wave function which must be orthogonal to the core electron wave function). It gives ample opportunity to introduce model potentials which simplify the calculations. The theory of atoms in molecules [5] is based on the idea that the relative ease with which atomic wave functions can be calculated can be used to construct molecular wave functions from these fragments. Related is the semiempirical method of diatomics in molecules [6] and the work of Adams on the solution of the Schrödinger and the HF equation in terms of wave functions which are least distorted from products of atomic wave functions [7–9]. The recent article of Ohno [10] *et al.* should also be mentioned in this context. The theory of the separability of many-electron systems in the case of the  $\pi$ -electron approximation has been examined among others by McWeeny [11] and by Lykos and Parr [12]. This problem has been in general treated by Huzinaga and Cantu [13]. The author has in a previous article presented a method which allows the construction of a wave function of a molecule starting from the wave functions of fragment molecules [14].

One of the most important concepts in the context of the present work is the concept of the localized molecular orbitals (LMO's) and their transferability property. LMO's closely correspond to the classical chemical concepts of inner shells and lone pairs of electrons and bonds linking the atoms in a molecule. They bridge the gap between these concepts and the rigorous description of molecules by wave functions. A wave function constructed from LMO's is equivalent to a wave function constructed from the delocalized canonical molecular orbitals, i.e. all expectation values of totally symmetric operators calculated with these wave functions are identical. LMO's can be obtained directly as solutions of the HF or multiconfiguration equations using a pseudopotential method [7, 15–19]. Peters has developed a different approach for the direct calculation of LMO's [20, 21]. These methods correspond to a reformulation of the HF equation, which can then be used to introduce approximations. (See e.g. Ref. [20]). But in most cases LMO's are obtained by a unitary transformation of the molecular orbitals (MO's) resulting as solutions of the HF equation in their standard form. LMO's have been introduced and examined mainly by Lennard-Jones and coworkers, by Boys, and by Edmister and Ruedenberg and they have proved to be extremely useful [22–29]. The aspect of the LMO's which is of greatest concern to the present work is their approximate transferability property. LMO's can be expected to be transferable among molecules having chemically related structures because they are maximally separated from each other and because they are themselves restricted to a minimal spatial region (for a discussion see Ref. [25]). Transferable orbitals have been the subject of a number of investigations [14, 24, 25, 29–37], although few calculations have been performed which calculate an atomic or molecular wave

function by an actual transfer of LMO's [14, 33, 34]. Shull *et al.* were to the knowledge of the author the first ones to report such a transfer of LMO's (in their case geminals) [33]. The detailed examination of the structure of the LMO's and an energy analysis based on them has also contributed to an understanding in this respect [24, 36]. The study of molecular momentum distributions and Compton profiles both by experimental and theoretical methods yields information on the transferability of LMO's and deserves notice in this context [37].

The transfer of LMO's or geminals and the transfer of parameters associated with them is only one aspect of the problem, although it constitutes the most frequently used path. Different approaches are possible. Lipscomb and coworkers made investigations in which they transferred matrix elements of the HF operator from smaller molecules to larger ones in order to calculate the wave function of the larger molecule [38]. Bader *et al.* defined a spatial partitioning of total molecular charge distributions independent of an orbital concept and discussed the transferability of molecular energies [39]. Christoffersen and coworkers transfer elements of the one-particle reduced density matrix between structurally related molecules, but use it only as a starting point for a SCF calculation [40]. Nelander discusses the partitioning of the first order density matrix and its use for bond energy schemes [41].

There are other constructive ways to generate molecular wave functions, which do not or which do not necessarily involve the variation principle. These methods start in general from the hybridization concept. The first step is the construction of hybrids which are combined to form one- or two-center localized bond orbitals using polarity parameters. These bond orbitals are then used to build the molecular wave function. Chemical experience and experimental data on the one hand or the variation principle on the other hand determine the necessary parameters. There are many variants of this method [42–44]. The wave function can surpass HF wave functions in quality, if the correlation energy is taken into account by perturbation theory [45]. The fundamental concept in these approaches is besides the concept of hybrids again the one of localized orbitals.

As mentioned above a method has been introduced which permits the computation of the wave function of a molecule from the wave functions of fragment molecules by transferring some of the LMO's of the fragments and recalculating the orbitals in the region of interaction [14]. In the present article this topic will be pursued and the method generalized. The formal theory will be presented in Sect. 2 together with a brief outline of the physical background. The application of the method to study the interaction of two Be atoms, the interaction of two  $\text{Li}_2$  molecules and to calculate the barrier to internal rotation in ethane is discussed in Sect. 3.

## 2. Theory

If one introduces a change in some part of any "large" molecule by a substitution or an isomerization, the effects of this change on the electronic distribution will be important mainly in its immediate neighbourhood. By

allowing for the electronic rearrangement in this immediate neighbourhood, which will be called the region of interaction, it should be possible to obtain a good approximation to the actual process in many cases. There will certainly be an electronic rearrangement in the far distant parts of the molecule too, but this should be only of minor importance for many practical questions. To take advantage of this one has to look for molecular fragments which allow to construct new molecules from these fragments such that the wave function of this new molecule constructed from the fragment wave functions is a good approximation to the wave function calculated by an *ab initio* method. For the general case of a molecule  $A-B$  being formed from two molecules  $A-X$  and  $B-Y$  (where  $A$ ,  $X$ ,  $B$ , and  $Y$  are any molecular fragments):

$$A-X+B-Y=A-B+\text{remainder}$$

one would distinguish three different spatial regions in each of the molecules  $A-X$  and  $B-Y$ : a region which will be discarded, a region of interaction, and a fixed core which will be transferred unaltered. This approach should open a general path to the calculation of approximate wave functions of molecules from molecular fragments.

The considerations will be restricted to the closed shell case of the wave functions constructed from real orbitals for all of the molecules  $A-X$ ,  $B-Y$ , and  $A-B$  [46]. Let

$$\begin{aligned} &\langle r_1, s_1, \dots, r_{2n_A}, s_{2n_A} | \Psi^A \rangle \\ &= (2n_A!)^{\frac{1}{2}} A^A \{ \langle r_1, s_1 | 1 + \rangle \langle r_2, s_2 | 1 - \rangle \dots \langle r_{2n_A}, s_{2n_A} | n_A - \rangle \} \end{aligned}$$

in short

$$|\Psi^A\rangle = (2n_A!)^{\frac{1}{2}} A^A \left\{ \prod_{i=1}^{n_A} |i+\rangle |i-\rangle \right\} \quad (1)$$

be the wave function for molecule  $A-X$  with  $2n_A$  electrons.  $A^A$  is the anti-symmetric projection operator for this case. The MO's  $|i\rangle$  are determined from the HF equations for  $A-X$

$$F^A |i\rangle = \sum_{i'} \varepsilon_{ii'} |i'\rangle, \quad (2)$$

where

$$F^A = h^A + \sum_i 2J_i^A - K_i^A. \quad (3)$$

$h^A(1) = -\frac{1}{2} \nabla_1^2 - \sum_A Z_A^A / R_{A1}$  and  $J_i^A$  and  $K_i^A$  are Roothaan's Coulomb and exchange operators [46]

$$\begin{aligned} J_i^A &= \langle i | r_{12}^{-1} | i \rangle \\ K_i^A &= \langle i | r_{12}^{-1} | i \rangle \langle i \rangle. \end{aligned} \quad (4)$$

$|+\rangle$  and  $|-\rangle$  denote the two possible spin states of a spin 1/2 system. The product is over all occupied MO's of molecule A - X. In the same way let

$$|\Psi^B\rangle = (2n_B!)^{\frac{1}{2}} A^B \left\{ \prod_{j=1}^{n_B} |j+\rangle |j-\rangle \right\} \quad (5)$$

be the wave function of molecule B - Y with similar definitions for the operator and quantities appearing in (5). The letter  $i$  will be used for the MO's of molecule A - X and the letter  $j$  for the MO's of molecule B - Y. The MO's  $|j\rangle$  are solutions of the HF equations for molecule B - Y:

$$F^B |j\rangle = \sum_j \epsilon_{jj'} |j'\rangle, \quad (6)$$

where

$$F^B = h^B + \sum_j 2J_j^B - K_j^B. \quad (7)$$

For a proper description of the different regions in the molecules it is necessary to construct the wave functions from LMO's. The orbitals  $|i\rangle$  and  $|j\rangle$  are thus the LMO's describing inner shells and lone pairs of electrons and bonds in the two molecules. LMO's can be defined as solutions of the HF equations corresponding to a non-diagonal matrix of Lagrangian multipliers [24]. This is why this form of the equations has been given in (2) and (6).

The following ansatz is made for the wave function of molecule A - B in this theory of molecules in molecules (MIM):

$$\begin{aligned} |\Psi\rangle &= (2n!)^{\frac{1}{2}} A \left\{ \left[ \prod_i' |i+\rangle |i-\rangle \right] \left[ \prod_j' |j+\rangle |j-\rangle \right] \left[ \prod_m |m+\rangle |m-\rangle \right] \right\} \\ &= (2n!)^{\frac{1}{2}} A \left\{ \left[ \prod_k |k+\rangle |k-\rangle \right] \left[ \prod_m |m+\rangle |m-\rangle \right] \right\} \end{aligned} \quad (8)$$

$A$  is the antisymmetric projection operator for the entire molecule containing  $2n$  electrons. Note that  $2n \neq 2n_A + 2n_B$  in general! The prime on the product sign indicates that a number of LMO's is left out to be deleted completely or to be recalculated in the region of interaction. The remaining LMO's (for which the letters  $k$  and  $l$  will be used) are transferred unaltered. They form the fixed core. These orbitals are nonorthogonal because they result from calculations on different molecules. The MO's  $|m\rangle$  are to be determined for the description of the new bonds formed and of their neighbourhood. They will be required to be orthogonal to the fixed orbitals  $|k\rangle$ . This can be done by using the projection operator for orthogonality

$$P = 1 - \sum_{k,l} |k\rangle S_{kl}^{-1} \langle l|; \quad P^2 = P, \quad P^+ = P, \quad (9)$$

where  $S^{-1}$  is the inverse matrix of the overlap matrix  $S = \{\langle k|l\rangle\}$  [47]. An outer projection by  $P$  of the operator determining the MO's  $|m\rangle$  gives orbitals exactly orthogonal to the MO's  $|k\rangle$  [47]. If the MO's  $|k\rangle$  can be regarded as

approximately orthogonal (e.g. because of spatial separation) this projection operator simplifies to

$$P = 1 - \sum_k |k\rangle \langle k|. \quad (10)$$

The approach suggested by Huzinaga and Cantu [13] is identical in the ansatz, and differs only by the use of the coupling operator technique [48, 49] instead of the present method of outer projections, if no further approximations are introduced.

The MO's  $|m\rangle$  are to be determined such that the total energy of the system is minimized. The total energy is calculated as the expectation value of the Hamiltonian operator  $H$  of the entire molecule A-B with the wave function of equation (8). The expression for the total energy in the case of nonorthogonal orbitals is given by

$$\begin{aligned} \langle \Psi | H | \Psi \rangle &= \sum_{n_1, n_2} \langle n_1 | h | n_2 \rangle D(n_1 | n_2) \\ &+ \frac{1}{2} \sum_{n_1, n_2, n_3, n_4} [n_1 n_2 | n_3 n_4] D(n_1 n_3 | n_2 n_4), \end{aligned} \quad (11)$$

where the  $|n_i\rangle$  are any MO's,

$$[ij|kl] := \iint \langle i | r_1 \rangle \langle r_1 | j \rangle r_{12}^{-1} \langle k | r_2 \rangle \langle r_2 | l \rangle d^3 r_1 d^3 r_2, \quad (12)$$

and  $D(i|j)$  and  $D(ij|kl)$  are minors of the determinant

$$D = \det \{ \langle i | j \rangle \}. \quad (13)$$

A clear discussion of MO theory for nonorthogonal orbitals can be found in a series of papers by Löwdin [50]. Variation of the orbitals  $|m\rangle$  to get the minimum energy in the restricted subspace created by the projection operator  $P$  leads to the equations

$$PFP|m\rangle = \epsilon_m |m\rangle, \quad (14)$$

where the matrix of Lagrangian multipliers is diagonal, because the MO's  $|m\rangle$  are not required to be LMO's. Note that the operators  $P$  and  $F$  in (14) do not commute because the MO's  $|k\rangle$  are not eigenfunctions of the operator  $F$ ! The HF operator  $F$  in (14) is given by

$$F = F_{\text{core}} + \sum_m 2J_m - K_m \quad (15)$$

$$F_{\text{core}} = h + \sum_{k,l} (2\langle k | r_{12}^{-1} | l \rangle - \langle k | r_{12}^{-1} | l \rangle \langle l | r_{12}^{-1} | k \rangle) S_{kl}^{-1}. \quad (16)$$

If the MO's  $|k\rangle$  are orthogonal or if their nonorthogonality is neglected, (16) simplifies to

$$F_{\text{core}} = h + \sum_k 2J_k - K_k. \quad (17)$$

where the familiar Coulomb and exchange operators appear. Since the projection is done prior to diagonalization optimal orthogonal MO's  $|m\rangle$  are obtained – optimal with respect to the fixed core.

The ansatz (8) and (14) represents an approximation to the exact HF wave function and HF equation. But the theory can be made exact by extending the work of Lykos and Parr [12] on the pi-electron approximation to the present case, where the orthogonality of the orbitals is not guaranteed by symmetry. (See also Ref. [13]). Above, the  $|m\rangle$ -electron approximation has been taken. If the MO's  $|k\rangle$  are now determined with the MO's  $|m\rangle$  held fixed, the  $|k\rangle$ -electron approximation is taken. This process can be continued until self-consistency is reached. The procedure is a generalization of the usual SCF procedure. The orthogonality of the orbitals is guaranteed by employing the method of outer projections, which corresponds to making the variations in a restricted subspace [47]. The MIM method can be made "exact" still in another way. If the region of interaction is the entire molecule, the wave function calculated is identical with the SCF wave function. This is again not the aim of the present approach, but it offers the possibility of obtaining wave functions differing in quality from the SCF wave function to more approximate ones. This method is thus capable of providing in a relatively nonarbitrary way an information on which MO's can be transferred, i.e. which MO's are not appreciably affected by the formation of the new bonds. Simultaneously it can give information on the energy contributions of the individual inner shells and lone pairs of electrons and bonds to the total energy change of a process.

The total SCF energy is always a lower limit to the energy expectation value in the MIM approximation, if the nonorthogonality of the MO's is properly taken into account. No general statement can be made if this nonorthogonality is neglected. But because the sets of transferred orbitals  $\{|i\rangle\}$  and  $\{|j\rangle\}$  are spatially well separated such an approximation should frequently be justified and give reliable answers.

The method discussed so far appears to be a reasonable approach which can be expected to give reliable answers. Applications made so far justify it [14, 51]. But it must be mentioned that the transfer of LMO's for some part of the molecule together with the redetermination of the MO's in the region of interaction does not lead to a considerable saving in the computation time. All integrals over the basis functions have to be evaluated, only the iteration part of the calculation might be shortened. This is not the final aim. A theory of molecules in molecules is desired which permits a significant saving in the computation time, i.e. gives the possibility to circumvent the " $N^4$  law" of the computational expense (where  $N$  is the number of particles or basis functions) at least in some part of the integral calculation. It is consequently necessary to proceed from the above ansatz and introduce further approximations.

Since only the MO's in the region of interaction are going to be redetermined and since the interest is in "larger" molecules, it might be acceptable to restrict the expansion of the MO's  $|m\rangle$  to a subset  $\Gamma$  of the total set of basis functions (a LCAO expansion form for the MO's is assumed from now on):

$$|m\rangle = \sum_{p \in \Gamma} |p\rangle C_{pm} \quad (18)$$

In mathematical terms this is an outer projection of the operator  $PFP$  of equation (14) by the projector on the manifold spanned by the basis functions in the set  $\Gamma$  according to

$$O_{\Gamma} P F P O_{\Gamma}, \quad (19)$$

where the projection operator  $O_{\Gamma}$  is given by

$$O_{\Gamma} = \sum_{p,q \in \Gamma} |p\rangle S_{pq}^{-1} \langle q|; O_{\Gamma}^2 = O_{\Gamma}, O_{\Gamma}^{\dagger} = O_{\Gamma}. \quad (20)$$

$S^{-1}$  is the inverse matrix of the overlap matrix  $S = \{\langle p|q\rangle\}$ ,  $p, q \in \Gamma$ . Such an approach reduces the dimension of the matrices to be diagonalized, but it does not affect the number of integrals which have to be evaluated. If this number is to be reduced as well, still another approximation must be made.

The basis functions whose centers are distant from the region of interaction should play only a minor role for the orthogonality of the MO's  $|m\rangle$  to the transferred LMO's  $|k\rangle$ . The LMO's  $|k\rangle$  to which orthogonality can be expected on spatial ground could be left out from the projection operator  $P$  and the basis functions which mainly contribute to these LMO's could be taken out from the entire set of basis functions for the expansion of the orbitals in the projection operator. This means the operator  $P$  in equation (9) is replaced by

$$P' = 1 - \sum'_{k,l} |k\rangle S_{kl}^{-1} \langle l|, \quad (21)$$

where the prime on the summation sign indicates the deletion of some of the orbitals. Then an outer projection of the operator  $P'$  is made according to

$$O_A P' O_A, \quad (22)$$

where

$$O_A = \sum_{r,s \in A} |r\rangle S_{rs}^{-1} \langle s|; O_A^2 = O_A, O_A^{\dagger} = O_A \quad (23)$$

is the projection operator on the manifold of the basis functions in the set  $A$ , which are to be included in the operator  $P'$ . This approximation results in a reduction of the number of integrals which have to be evaluated. For the success of the method it is essential that this approximation works. The final operator determining the MO's  $|m\rangle$  is given by

$$O_{\Gamma} O_A P' O_A F O_A P' O_A O_{\Gamma} \quad (24a)$$

Note:  $I'$  must always be a subset of  $A: I' \subseteq A$ ! The other case is physically meaningless. The operator therefore simplifies to

$$O_{\Gamma} P' O_A F O_A P' O_{\Gamma} \quad (24b)$$

because

$$O_{\Gamma} O_A = O_A O_{\Gamma} = O_{\Gamma}. \quad (25)$$

This operator permits to save computational time, but it has acquired a relatively complicated structure. One main disadvantage is that the additional approximations introduced by the projections result in a nonorthogonality of the MO's



$|m\rangle$  to the transferred LMO's  $|k\rangle$ . The final justification of the approach can only be done by calculations.

Although these equations have been derived for the case of two fragments joined together, it is easily extended to any number of fragments. In fact the final equations are valid for the general case.

A few more points have to be discussed in this context. If one wants to avoid the calculation of all integrals, which is one intention of the theory of molecules in molecules, the nonorthogonality of the MO's has to be neglected – otherwise all integrals will appear again. The expression for the electronic energy then takes the well-known form

$$E_{el} = 2 \sum_n \langle n|h|n \rangle + \sum_{n_1, n_2} 2J_{n_1 n_2} - K_{n_1 n_2} \quad (26)$$

where  $J_{n_1 n_2}$  and  $K_{n_1 n_2}$  are the Coulomb and exchange integrals between any orbitals  $|n_1\rangle$  and  $|n_2\rangle$ . In this expression the interaction terms  $J_{ij}$  and  $K_{ij}$  between an orbital transferred from molecule A – X and another orbital transferred from molecule B – Y appear. These terms necessitate the calculation of the integrals over all basis functions, if one is not willing to approximate some of these integrals or to approximate the energy terms  $J_{ij}$  and  $K_{ij}$  directly. The simplest method would be to make a point charge approximation for every orbital, which would be justified for large distances:

$$\begin{aligned} K_{ij} &= 0 \\ J_{ij} &= |r_i - r_j|^{-1}, \end{aligned} \quad (27)$$

where  $r_i = \langle i|R|i \rangle$  is the charge centroid of LMO  $|i\rangle$ .

This particular approximation is not necessary and more refined procedures can be used. All two-electron integrals have then to be evaluated exactly except the integrals between the basis functions in the two fixed cores; this will be an appreciable part for larger molecules.

Equation (24) for the operator defining the MO's  $|m\rangle$  is complicated and because of the multiple projection the question arises, whether it is still possible to give bounds to the eigenvalues and energy expectation values as in the case of a simple projection. The HF operator is projected according to  $Q^+ F Q$  by an operator  $Q = O_A P' O_I \cdot Q$  is a product of projection operators, but itself is not a projection operator due to the fact that the projection operators  $O_I$  and  $O_A$  do not commute with  $P'$  (i.e.  $Q^2 \neq Q$  and  $Q^+ \neq Q$ ). However, it is still possible to derive bounds. Löwdin [47] has shown that for an outer projection of a self-adjoint operator  $A$ , which is bounded from below, by an arbitrary projection operator  $O$

$$O A O = \bar{A} \quad \text{with} \quad O^2 = 0, O^+ = 0 \quad (28)$$

the eigenvalues of  $\bar{A}$  are upper bounds to the eigenvalues of  $A$  in order

$$a_k \leq \bar{a}_k, \quad (29)$$

where

$$A|k\rangle = a_k|k\rangle \quad \text{and} \quad \bar{A}|\bar{k}\rangle = \bar{a}_k|\bar{k}\rangle. \quad (30)$$

$|k\rangle$  and  $|\bar{k}\rangle$  are normalized eigenfunctions. (See also Ref. [52]). Let  $P$  be another projection operator

$$P^2 = P, P^+ = P \quad (31)$$

which does not commute with  $O$

$$[P, O] \neq 0. \quad (32)$$

According to the definition the operator

$$P\bar{A}P = \bar{A} \quad (33)$$

is an outer projection of  $\bar{A}$  with respect to  $P$ . The eigenvalue equation for  $\bar{A}$  is

$$\bar{A}|\bar{k}\rangle = \bar{a}_k|\bar{k}\rangle \quad (34)$$

and applying Löwdin's theorem again the eigenvalues  $\bar{a}_k$  are upper bounds to the eigenvalues of  $\bar{A}$  in order. Using equation (29) one obtains

$$a_k \leq \bar{a}_k \leq \bar{a}_k. \quad (35)$$

This result applies in the following form to the projected HF operator of the theory of molecules in molecules:

$$\begin{aligned} e_m(F) &\leq e_m(O_A F O_A) \leq e_m(P' O_A F O_A P') \leq e_m(O_A P' O_A F O_A P' O_A) \\ &\leq e_m(O_I O_A P' O_A F O_A P' O_A O_I), \end{aligned} \quad (36)$$

where the  $e_m$  are the eigenvalues of the operators given in brackets. The same series of inequalities can be written down for the energy expectation value. (The operators are given in brackets.)

$$\begin{aligned} E(F) &\leq E(O_A F O_A) \leq E(P' O_A F O_A P') \leq E(O_A P' O_A F O_A P' O_A) \\ &\leq E(O_I O_A P' O_A F O_A P' O_A O_I). \end{aligned} \quad (37)$$

This result will be taken up again in the applications.

Another important bound can be derived in a similar way as above. Let the sets  $F'$  and/or  $A'$  be obtained from  $F$  and/or  $A$  by subtracting some functions (subject to the restriction that  $F' \subseteq A'$  whatever members the two sets have). For the projection operators  $O_{F'}$  and  $O_{A'}$  the following relations hold

$$O_{F'} O_I = O_I O_{F'} = O_{F'} \quad (38)$$

and

$$O_{A'} O_A = O_A O_{A'} = O_{A'}. \quad (39)$$

The reduction of the number of basis functions can thus be written as another outer projection

$$O_{F'} O_I P' O_A F O_A P' O_I O_{F'} = O_{F'} P' O_A F O_A P' O_{F'} \quad (40)$$

and

$$O_I P' O_{A'} O_A F O_A O_{A'} P' O_I = O_I P' O_{A'} F O_{A'} P' O_I. \quad (41)$$

Because of this fact one can immediately write down the following inequalities:

$$\varepsilon_m(O_T P' O_A F O_A P' O_T) \leq \varepsilon_m(O_T P' O_A F O_A P' O_T) \quad (42)$$

$$\varepsilon_m(O_T P' O_A F O_A P' O_T) \leq \varepsilon_m(O_T P' O_A F O_A P' O_T) \quad (43)$$

and for the energy expectation value

$$E(O_T P' O_A F O_A P' O_T) \leq E(O_T P' O_A F O_A P' O_T) \quad (44)$$

$$E(O_T P' O_A F O_A P' O_T) \leq E(O_T P' O_A F O_A P' O_T) \quad (45)$$

This supplies the rigorous basis for the statement made before concerning the quality of the approximated wave functions. The physical argument is that one is working in more and more restricted subspaces which consequently raises the energy.

In the applications the justification of the four approximations introduced by 1)  $O_T$ , 2)  $O_A$  together with the approximation of  $P$  by  $P'$ , 3) the neglect of the nonorthogonality of the MO's, and 4) the point charge approximation in the calculation of the total energy will be examined.

### 3. Applications

#### 3.A. Be Be

Wave functions have been calculated for the  $\text{Be}_2$  molecule for various distances in the range from 2 to 12 a.u. On each atom a basis set of 9  $s$ -type Gaussian functions [53], which are left completely uncontracted, is used to expand the MO's. For the Be atom the resulting total SCF energy is  $E^{\text{SCF}} = -14.572068$  a.u. This compares with the value calculated by Clementi using a Slater-type basis:  $E^{\text{SCF}} = -14.573020$  a.u. [54]. The energy values calculated for the  $\text{Be}_2$  molecule by the *ab initio* method are given in Table 1, the potential curve is a repulsive curve. The calculated wave functions serve as a simple laboratory to test the approximations introduced in Sect. 2. The canonical MO's are transformed to LMO's by the method of Edmiston and Rueden-

Table 1. Total SCF energies for the  $\text{Be}_2$  molecule (all values in atomic units)

$R$	$E^{\text{SCF}}$
2.0	-28.457626
3.0	-28.937517
4.0	-29.072896
6.0	-29.134913
8.0	-29.143207
12.0	-29.144131
$\infty$	-29.144137

Table 2. Total energies for the  $\text{Be}_2$  molecule calculated in the MIM approximation (all values in atomic units, for notation see text)

$R$	$E^{\text{MIM}}_a$	$E^{\text{MIM}}_b$	$E^{\text{MIM}}_a$	$E^{\text{MIM}}_b$
Approximation $\Gamma$ 18 $\Delta$ 18		Approximation $\Gamma$ 16 $\Delta$ 18		
2.0	28.450985	-28.451053	-28.450818	-28.453928
3.0	28.932477	-28.934408	-28.932356	-28.937243
4.0	29.070994	-29.071620	-29.070897	-29.074349
6.0	29.134804	-29.134818	-29.134730	-29.137385
8.0	29.143202	-29.143195	-29.143135	-29.145732
12.0	29.144131	29.144131	-29.144066	-29.146664
Approximation $\Gamma$ 14 $\Delta$ 18		Approximation $\Gamma$ 12 $\Delta$ 18		
2.0	28.449401	-28.480178	-28.446801	-28.577635
3.0	28.931497	-28.960672	-28.929688	-29.050076
4.0	29.070222	-29.096477	-29.068847	-29.179064
6.0	29.134170	-29.158525	-29.133064	-29.232825
8.0	29.142623	-29.166574	-29.141584	-29.238986
12.0	29.143565	29.167444	29.142538	-29.239556
Approximation $\Gamma$ 16 $\Delta$ 16		Approximation $\Gamma$ 14 $\Delta$ 16		
2.0	28.450814	-28.450853	-28.449345	-28.473663
3.0	28.932354	-28.934262	-28.931457	-28.954868
4.0	29.070895	-29.071501	-29.070189	-29.091052
6.0	29.134728	-29.134723	-29.134145	-29.153459
8.0	29.141133	-29.143108	-29.142602	-29.161601
12.0	29.144064	29.144045	-29.143544	-29.162488
Approximation $\Gamma$ 12 $\Delta$ 16		Approximation $\Gamma$ 14 $\Delta$ 14		
2.0	28.446708	-28.573309	-28.449091	-28.448388
3.0	28.929626	-29.046714	-28.931284	-28.932593
4.0	29.068796	-29.176071	-29.070045	-29.070129
6.0	29.133022	-29.230020	-29.134037	-29.133534
8.0	29.141545	-29.236234	-29.142508	-29.141982
12.0	29.142500	-29.236815	-29.143453	-29.142933
Approximation $\Gamma$ 12 $\Delta$ 14		Approximation $\Gamma$ 12 $\Delta$ 12		
2.0	-28.445775	-28.534407	-28.441889	-28.434050
3.0	-28.928968	-29.014654	-28.925911	-28.921521
4.0	-29.068250	-29.147044	-29.065682	-29.060888
6.0	29.132577	29.203173	-29.130656	-29.125515
8.0	29.141140	-29.209981	-29.139443	-29.134258
12.0	29.142104	29.210677	-29.140453	-29.135264

a nonorthogonality of the MO's taken into account

b nonorthogonality neglected

berg [24]. The approximations are applied in the following way. The inner shell LMO's will be transferred and the outer shell orbitals recalculated. A series of calculations has been made with different numbers of basis functions in the sets  $\Gamma$  and  $\Delta$ . The following notation will be used to define the individual calculations. For both sets  $\Gamma$  and  $\Delta$  the total number of functions which they

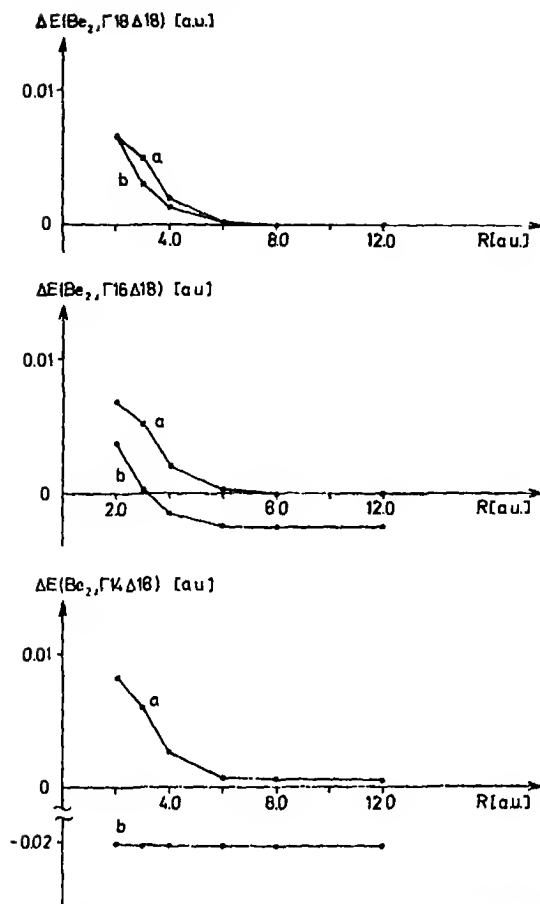


Fig. 1. Energy difference between MIM and SCF result for  $\text{Be}_2$ :  $\Delta E = E^{\text{MIM}} - E^{\text{SCF}}$ ; approximations  $\Gamma 18 \Delta 18$ ,  $\Gamma 16 \Delta 18$ ,  $\Gamma 14 \Delta 18$  (for notation see text). a: nonorthogonality of the MO's taken into account, b: nonorthogonality neglected

contain is written behind the respective symbol.  $\Gamma 18 \Delta 18$  e.g. means that all basis functions are included. For the other approximations functions are excluded stepwise from the two sets, where the deletion is done in the sequence of decreasing exponential parameter of the functions. The calculations performed are defined by the symbols  $\Gamma 18 \Delta 18$ ,  $\Gamma 16 \Delta 18$ ,  $\Gamma 14 \Delta 18$ ,  $\Gamma 12 \Delta 18$ ,  $\Gamma 16 \Delta 16$ , ... and finally  $\Gamma 12 \Delta 12$ . In addition the letter *a* denotes that the energy expectation value is calculated taking the nonorthogonality of the MO's into account, the letter *b* denotes the neglect of this nonorthogonality. For the orthogonalization the method of Löwdin has been used [55]. The total energies, which have been computed for all the approximations, are given in Table 2. In Fig. 1 the energy differences  $\Delta E = E^{\text{MIM}} - E^{\text{SCF}}$  are plotted for some examples which are representative for the whole set of calculations: the curve for  $\Gamma 12 \Delta 18$  is similar to the one for  $\Gamma 14 \Delta 18$ ;  $\Gamma 16 \Delta 16$  corresponds to  $\Gamma 18 \Delta 18$ , from which it differs

Table 3. Interaction energy between the two sets of transferred LMO's for  $\text{Be}_2$  calculated exactly ( $X_{ex}$ ) and by a point charge approximation ( $X_{pc}$ ). The nonorthogonality of the MO's is neglected. (All values in atomic units)

$R$	$X_{ex}$	$X_{pc}$	$\Delta X = X_{pc} - X_{ex}$
2.0	1.989410	2.0	0.010590
3.0	1.330469	1.333333	0.002864
4.0	0.999347	1.0	0.000653
6.0	0.666638	0.666667	0.000029
8.0	0.499999	0.5	0.000001
12.0	0.333333	0.333333	0.0

very little;  $F14A16$  is similar to  $F14A18$  etc. Considering the figure the following statements can be made. The curves  $a$  and  $b$  do not meet in general for  $R \rightarrow \infty$  and do in general not approach the limiting value zero, but they run parallel to the abscissa for interatomic distances larger than approximately 6 a.u. (for some cases the poorer approximations  $\sim 8$  a.u.). This then corresponds to a parallel shift of the potential curves, which is the result one would like to obtain. If a number of basis functions is completely excluded from the expansion of the outer shell MO's which is the case for most of the approximations — the limiting value zero for the energy difference  $\Delta E$  can no longer be reached and further on the curves  $a$  and  $b$  can no longer meet. It is the nonorthogonality of the outer shell MO on any Be atom to the inner shell MO on the same atom which causes this. The first conclusion to be drawn is that it is advisable to keep the sets  $I$  and  $A$  identical, which will be done in the subsequent investigations. The difference between curves  $a$  and  $b$ , which determines the reliability of the calculations where the nonorthogonality is neglected in the computation of the energy expectation value, is minimal for this case. There is in fact little computational advantage in allowing the sets  $I$  and  $A$  to differ. The operator defining the MO's  $|m\rangle$  thus simplifies to ( $O_I = O_A = O$ ):

$$O P' O F O P' O. \quad (46)$$

In Table 3 the energy values for the exactly calculated interaction energy of the transferred LMO's, for the point charge approximation to it, and their difference  $\Delta X = X_{pc} - X_{ex}$  is given. This approximation seems to be justified for distances greater than about 6 a.u.

### 3.B. $\text{Li}_2$ - $\text{Li}_2$

The interaction of two  $\text{Li}_2$  molecules is more interesting than the interaction of two Be atoms. 9  $s$ -type functions contracted to 6  $s$ -type functions [53] augmented by 2  $p\sigma$ -type Gaussian lobe functions were used on every Li atom. The parameters for the  $p$ -type functions are:  $\eta_1 = 0.5$ ,  $R_1 = \pm 0.075$  a.u. and  $\eta_2 = 2.0$ ,  $R_2 = \pm 0.065$  a.u. For the experimental bond length  $R = 5.0504$  a.u. the calculated total SCF energy is  $E^{\text{SCF}} = -14.862665$  a.u. This compares with the result of

Table 4. Total SCF energies for the linear geometry of the  $\text{Li}_4$  molecule (the bond distance of the two  $\text{Li}_2$  molecules is kept fixed at its experimental value of 5.0504 a.u.; all values in atomic units)

$R$	$E^{\text{SCF}}$
3.0	-29.645618
4.0	-29.699425
5.0504	-29.720226
6.0	-29.725574
8.0	-29.726184
12.0	-29.725281
20.0	-29.725316
$\infty$	-29.725330

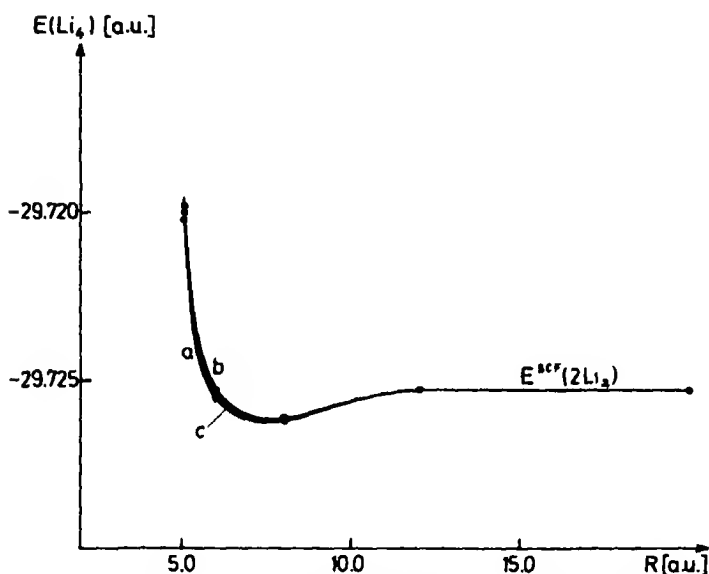


Fig. 2. Potential energy curve for  $\text{Li}_2 - \text{Li}_2$ ; approximation 1, 1, 4 / 32 (for notation and definition of a, b, c see text)

Ransil and Sinai:  $E^{\text{SCF}} = -14.87152$  a.u. [56]. The bond distance of the two  $\text{Li}_2$  molecules is held fixed at  $R = 5.0504$  a.u., only the distance between the two molecules is varied for the linear configuration in the range from  $R_{\text{LiLi}'} = 3.0$  a.u. to  $R_{\text{LiLi}'} = 20.0$  a.u. The energy values are given in Table 4 and the potential curve is plotted in Fig. 2, curve a. It is interesting to observe that the  $\text{Li}_4$  molecule is bound relative to two  $\text{Li}_2$  molecules with a binding energy of approximately 0.54 kcal/mole at a distance of about 7.5 a.u. At a distance of about 12 a.u. there is an extremely small maximum of approximately 0.03 kcal/mole in the potential curve, but this value is too small to be interpretable. The MO's have been localized by the method of Edmiston and Ruedenberg [24].

Table 5. Total Energies for the linear geometry of the  $\text{Li}_4$  molecule calculated in the MIM approximation. The energy values are given for the exact calculation ( $E_{ex}^{\text{MIM}}$ ) and for the calculation involving the point charge approximation ( $E_{pc}^{\text{MIM}}$ ) both for the neglect of the nonorthogonality ( $b$ ) and taking it into account ( $a$ ). (All values in atomic units, for notation see text)\*

$R$	$E_{ex}^{\text{MIM}}(a)$	$E_{ex}^{\text{MIM}}(b)$	$E_{pc}^{\text{MIM}}(a)$	$E_{pc}^{\text{MIM}}(b)$
Approximation 1, 1, 4 $\Gamma$ 32				
3.0	29.644888	- 29.644876	- 29.645150	- 29.645136
4.0	- 29.698830	- 29.698818	- 29.699034	- 29.699021
5.0504	29.719834	- 29.719825	- 29.719994	- 29.719984
6.0	29.725334	- 29.725328	- 29.725464	- 29.725457
8.0	29.726116	- 29.726114	- 29.726203	- 29.726201
12.0	29.725278	29.725278	- 29.725323	- 29.725324
20.0	29.725316	- 29.725316	- 29.725333	- 29.725333
Approximation 1, 1, 4 $\Gamma$ 30				
3.0	29.638803	29.635818	- 29.639110	- 29.636078
4.0	29.693411	29.690324	- 29.693644	- 29.690526
5.0504	- 29.714923	- 29.711548	- 29.715098	- 29.711706
6.0	29.720727	- 29.717065	- 29.720863	- 29.717194
8.0	29.721801	- 29.717732	- 29.721888	- 29.717819
12.0	29.721022	29.716755	- 29.721064	- 29.716800
20.0	29.721037	- 29.716749	- 29.721053	- 29.716766
Approximation 2, 2, 2 $\Gamma$ 30				
3.0	29.636990	29.634395	- 29.643269	- 29.640672
4.0	29.692931	29.689996	- 29.696687	- 29.693735
5.0504	29.714643	- 29.711327	- 29.717108	- 29.713778
6.0	29.720555	29.716899	- 29.722317	- 29.718656
8.0	29.721774	- 29.717687	- 29.722721	- 29.718642
12.0	29.721059	29.716779	- 29.721426	- 29.717153
20.0	- 29.721078	- 29.716778	- 29.721182	- 29.716884
Approximation 2, 2, 2 $\Gamma$ 28				
3.0	29.629013	29.619032	- 29.634952	- 29.625309
4.0	- 29.686023	29.676296	- 29.689619	- 29.680035
5.0504	- 29.709180	- 29.700053	- 29.711548	- 29.702504
6.0	29.715760	- 29.706846	- 29.717455	- 29.708603
8.0	29.717376	- 29.708449	- 29.718291	- 29.709404
12.0	29.716747	- 29.707682	29.717102	- 29.708056
20.0	29.716796	- 29.707719	- 29.716896	- 29.707826
Approximation 3, 3, 0 $\Gamma$ 32				
3.0	- 29.623096	- 29.800562	- 29.629370	- 29.797858
4.0	- 29.687252	- 29.787452	- 29.739871	- 29.847213
5.0504	- 29.713562	- 29.762680	- 29.778986	- 29.835688
6.0	29.721609	- 29.746389	- 29.784926	- 29.814538
8.0	- 29.724955	- 29.730399	- 29.772598	- 29.778947
12.0	- 29.725239	- 29.725417	- 29.747492	- 29.747632
20.0	- 29.725316	- 29.725316	- 29.731914	- 29.731914

\*The bond distance of the two  $\text{Li}_2$  molecules is kept fixed at its experimental value of  $R = 5.0504$  a.u.



A number of calculations have been made applying the method of molecules in molecules in various approximations. They will be described in the sequence of decreasing accuracy. The inner shell MO's on the two outer Li atoms are transferred and the set  $\Gamma$  includes all basis functions (i.e. altogether 32) in the first application. This approximation is denoted as 1, 1, 4  $\Gamma$  32; the notation gives the number of LMO's transferred for the first and the second  $\text{Li}_2$  molecule and the number of MO's redetermined in the region of interaction (4). By excluding the contracted  $s$ -type function on both outer Li atoms approximation 1, 1, 4  $\Gamma$  30 results. If the inner shell MO's on the two inner Li atoms are transferred as well, the approximations 2, 2, 2  $\Gamma$  30 and 2, 2, 2  $\Gamma$  28 are obtained, where  $\Gamma$  28 means that the contracted  $s$ -type function is excluded on all four Li atoms from the expansion of the bonding orbitals. In the crudest approximation all LMO's are transferred (3, 3, 0  $\Gamma$  32). The energy values for all approximations are listed in Table 5. The potential curves for approximation 1, 1, 4  $\Gamma$  32 are plotted in Fig. 2, curves  $b$  and  $c$  ( $b$ : interaction energy between the transferred LMO's calculated exactly,  $c$ : interaction energy calculated by the point charge approximation). Both curves  $b$  and  $c$  are in excellent agreement with the potential curve  $a$  calculated by the *ab initio* SCF method. The nonorthogonality of the MO's causes nearly no error in this case as can be seen from Table 5. The corresponding potential curves coincide with curves  $b$  and  $c$ , respectively. The difference

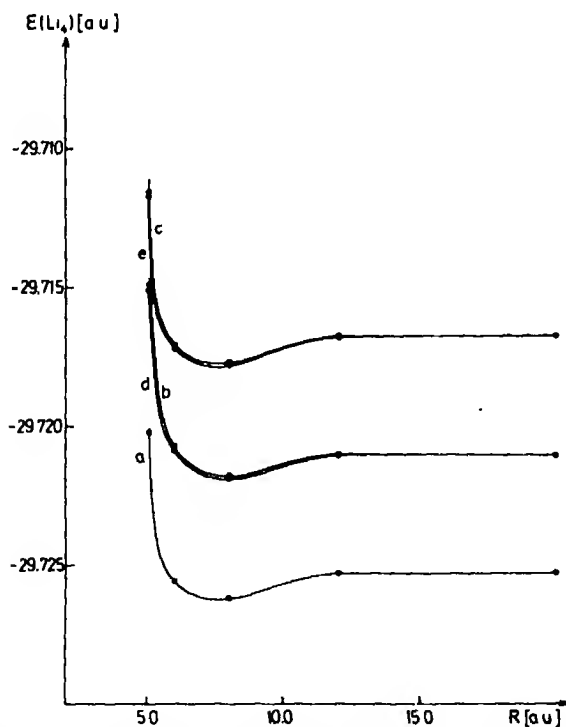


Fig. 3. Potential energy curve for  $\text{Li}_2 - \text{Li}_2$ ; approximation 1, 1, 4  $\Gamma$  30 (for notation and definition of  $a, b, c, d, e$  see text)

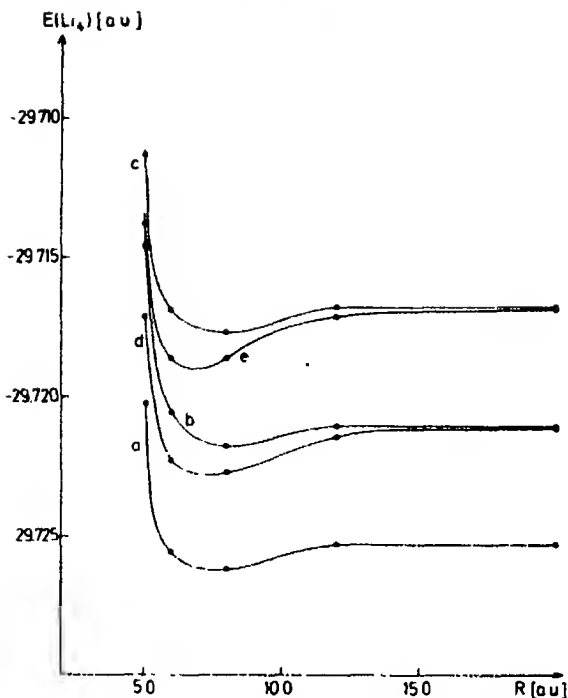


Fig. 4 Potential energy curve for  $\text{Li}_2 \cdot \text{Li}_2$ ; approximation 2, 2, 2  $\Gamma$  30 (for notation and definition of  $a, b, c, d, e$  see text)

$\Delta E = E^{\text{MIM}} - E^{\text{SCF}}$  is nearly zero for all  $R$ . For approximation 1, 1, 4  $\Gamma$  30 the potential curves are plotted in Fig. 3 (curve  $a$ : SCF result,  $b$ : energy value calculated exactly, nonorthogonality of the MO's taken into account [55],  $c$ : as  $b$  but nonorthogonality neglected,  $d$ : energy value calculated including point charge approximation, nonorthogonality taken into account [55],  $e$ : as  $d$  only with nonorthogonality neglected). Curves  $b$  and  $c$  are nearly parallelly shifted from the exact SCF curve and curves  $d$  and  $e$  are good approximations to them.

The agreement deteriorates slightly for approximation 2, 2, 2  $\Gamma$  30 (Table 5 and Fig. 4, the letters  $a, b, c, d$  and  $e$  have the same meaning as above). Curves  $b$  and  $c$  are still fairly parallel to  $a$ . But the point charge approximation in the calculation of the energy leads to a small shifting of the minimum to smaller  $R$  values and to a deeper minimum. This tendency is found as well for approximation 2, 2, 2  $\Gamma$  28 (Table 5 and Fig. 5 with  $a, b, c, d, e$  as defined above). For a good agreement with the SCF result (parallel shift of the potential curve)  $R$  must be larger than about 6–8 a.u.

The approximation 3, 3, 0  $\Gamma$  32, in which all MO's have been transferred, is unreasonable because the minimum in the potential curve is not reproduced (Table 5). It is concluded that it is the modification of the bonding orbitals in the two  $\text{Li}_2$  molecules which leads to bonding in  $\text{Li}_4$ .

Table 6 summarizes the effect of the point charge approximation on the interaction energy between the nonorthogonal transferred MO's:  $a$  is a very good

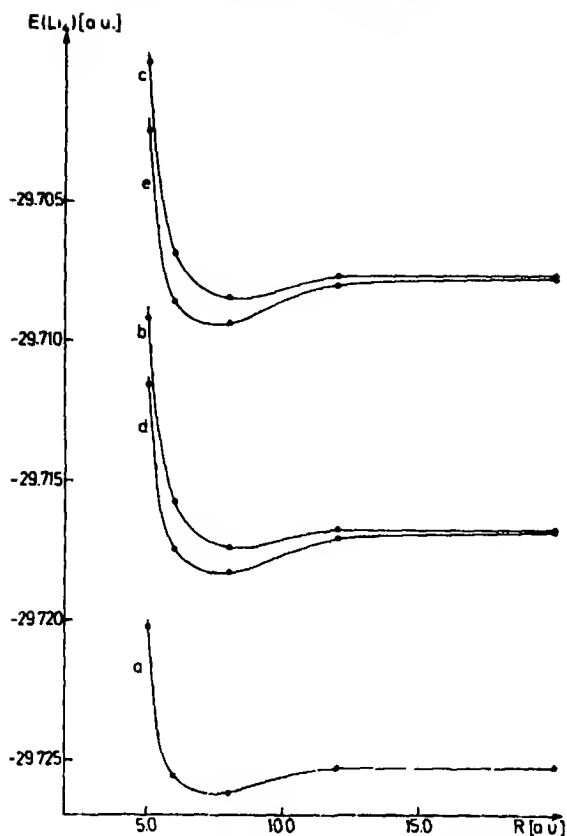


Fig. 5. Potential energy curve for  $\text{Li}_2 - \text{Li}_2$ ; approximation 2, 2, 2  $\Gamma$  28 (for notation and definition of  $a, b, c, d, e$  see text)

Table 6. Difference between the exactly calculated interaction energy between the two sets of transferred LMO's and the result of the point charge approximation for  $\text{Li}_4$  ( $\Delta X = X_{pc} - X_{ex}$ ). The nonorthogonality of the MO's is neglected. (All values in atomic units; for notation see text)

$R$	a $\Delta X$ (1, 1, 4 $\Gamma$ 32 and 1, 1, 4 $\Gamma$ 30)	b $\Delta X$ (2, 2, 2 $\Gamma$ 30 and 2, 2, 2 $\Gamma$ 28)	c $\Delta X$ (3, 3, 0 $\Gamma$ 32)
3.0	-0.000260	-0.006277	0.002704
4.0	-0.000202	-0.003739	-0.059761
5.0504	-0.000158	-0.002451	-0.073009
6.0	-0.000129	-0.001757	-0.068149
8.0	-0.000087	-0.000955	-0.048548
12.0	-0.000045	-0.000374	-0.022214
20.0	-0.000017	-0.000106	-0.006599

Table 7 Approximate bond distance and binding energy  $B$  of the  $\text{Li}_4$  molecule relative to two  $\text{Li}_2$  molecules (kept at their experimental bond length of  $R = 5.0504$  a.u.). (For an explanation of the notation see the text)

Method	$R_{\text{LiLi}}$ [a.u.]	$B$ [kcal/mole]
SCF	7.5	0.54
1, 1, 4 / 32 b	7.5	0.5
1, 1, 4 / 32 c	7.5	0.55
1, 1, 4 / 30 b	7.5	0.49
1, 1, 4 / 30 c	7.5	0.62
1, 1, 4 / 30 d	7.5	0.52
1, 1, 4 / 30 e	7.5	0.66
2, 2, 2 / 30 b	7.5 - 8.0	0.44
2, 2, 2 / 30 c	7.5 - 8.0	0.57
2, 2, 2 / 30 d	7.0	1.0
2, 2, 2 / 30 e	7.0	1.1
2, 2, 2 / 28 b	8.0	0.36
2, 2, 2 / 28 c	8.0	0.46
2, 2, 2 / 28 d	7.0 7.5	0.88
2, 2, 2 / 28 e	7.0 7.5	0.99
3, 3, 0 / 32	$\infty$	0.0

result,  $b$  is satisfactory for  $R > 8.0$  a.u. and  $c$  is completely useless, but this is not surprising.

In Table 7 the approximate bond distance  $R_{\text{LiLi}}$  and binding energy of the  $\text{Li}_4$  molecule with respect to two  $\text{Li}_2$  molecules are given as a summary on the quality of the various approximations investigated.

### 3.C. Barrier to Internal Rotation in Ethane

The basis set used in the calculation of the barrier to internal rotation in ethane consists of 7  $s$ -type functions and 3  $p$ -type Gaussian lobe functions on every C atom [57] contracted to 5  $s$ -type and 2  $p$ -type functions and of 3  $s$ -type functions on every H atom contracted to 2  $s$ -type functions [53]. For the  $p$ -type functions the distances from the center have been chosen to be  $R_1 = \pm 0.1$  a.u. for  $\eta_1 = 0.1992$ ,  $R_2 = \pm 0.08$  a.u. for  $\eta_2 = 0.8516$  and  $R_3 = \pm 0.05$  a.u. for  $\eta_3 = 4.1829$ . The experimental geometry has been used for both the staggered and the eclipsed conformation of the ethane molecule. The total energy calculated for the staggered form is  $E^{\text{SCF}} = -79.090587$  a.u. and for the eclipsed form  $E^{\text{SCF}} = -79.085792$  a.u., which results in a rotational barrier of 3.01 kcal/mole. The best wave function for  $\text{C}_2\text{H}_6$  has been calculated by Veillard, who obtained  $E^{\text{SCF}} = -79.2377$  a.u. for the staggered form and a rotational barrier of 3.07 kcal/mole [58]. The experimental value is 2.928 kcal/mole [59].

The theory of molecules in molecules is applied in the following way. Two  $\text{CH}_4$  molecules in the appropriate geometry serve as fragments for the  $\text{C}_2\text{H}_6$  molecule. One of the H atoms and its associated C-H bond orbital is taken

out from each of the two  $\text{CH}_4$  molecules. The inner shell MO's and the remaining 6 C-H bond orbitals are transferred. Thus only the C-C bond orbital has to be determined. The different approximations, which have been investigated, are described as follows: 1)  $\Gamma 34$ : all basis functions are included in  $\Gamma$ , 2)  $\Gamma 28$ : the contracted s-type function on all of the H atoms is taken out, 3)  $\Gamma 22$ : all basis functions on the H atoms are taken out, 4)  $\Gamma 20$ : in addition to 3) the contracted s-type functions used mainly to describe the inner shell MO's on the C atoms are taken out. The exact calculation of the energy with the nonorthogonality of the MO's taken into account is denoted by  $a$  [55], if the nonorthogonality is neglected,  $b$  is obtained;  $c$  and  $d$  are the corresponding results when the interaction energy between the two sets of LMO's transferred from the fragments is calculated by the point charge approximation. The transfer has been done for the energy localized MO's of Edmiston and Ruedenberg [24] (ELMO), the LMO's of Boys [23], the LMO's of Magnasco and Perico [26] and the density localized MO's of the author [29]. It turns out that the results for the different LMO's are nearly the same throughout in agreement with the fact that the LMO's themselves do not differ significantly [29]. Only the data for the ELMO's will be given. It further on turns out that the approximations denoted by  $\Gamma 34$ ,  $\Gamma 28$ ,  $\Gamma 22$ , and  $\Gamma 20$  all give nearly the same value for the rotational barrier and it is only of importance whether one is dealing with case  $a$ ,  $b$ ,  $c$  or  $d$  of the calculations. The results for the rotational barrier are (in kcal/mole):  $\Delta E(a) = 2.38$ ,  $\Delta E(b) = -2.06$ ,  $\Delta E(c) = 3.34$ , and  $\Delta E(d) = 1.52$ . (The values are given for approximation  $\Gamma 34$ , the values for the other cases differ by less than 0.1 kcal/mole). It is seen that the nonorthogonality of the MO's is important causing result  $b$  to have the incorrect sign. If the nonorthogonality is properly taken into account, the rotational barrier is about 20% smaller than the *ab initio* value, which is satisfactory. The point charge approximation gives an acceptable result in both cases and result  $c$  is quite a good approximation to  $a$ , which is surprising if one considers how unjustified this point charge approximation in the energy evaluation looks for such a small molecule as ethane. The fact that the other approximations (indicated by the notation  $\Gamma 28$ ,  $\Gamma 22$ ,  $\Gamma 20$ ) do not remarkably affect the value of the rotational barrier is somewhat surprising in view of the crudeness of some of the approximations made. Karplus and coworkers [60] have analyzed the rotational barrier in ethane. They were able to show that nearly any wave function gives an acceptable value for the barrier, if it is only antisymmetric. From the present investigation one can in addition conclude that the antisymmetry with respect to the two  $\text{CH}_3$  fragments does not play a crucial role, because this has been neglected in the calculation of  $E^{\text{MIM}}$  by the point charge approximation, but only the antisymmetry within each  $\text{CH}_3$  group taken together with the C-C bond orbital.

The approximations made for  $\text{C}_2\text{H}_6$  can lead to a considerable saving in the computation time. For the approximation denoted by  $\Gamma 28$  a saving of about 47% in the integral computation could be achieved, for  $\Gamma 22$  and  $\Gamma 20$  a saving of about 59% and 60%, respectively, would be possible. These numbers are given only as an orientation. They have to be taken with great care, because many integrals would anyway be zero or negligible and secondly the necessity of taking the non-orthogonality of the MO's into account modifies this result.

#### 4. Conclusions

The following conclusions can be drawn from these applications. The construction of the wave function of a molecule from the wave functions of fragment molecules by transferring some of the LMO's and redetermining the MO's in the region of interaction is a justified approach which gives results in close agreement with those obtained by the *ab initio* method. In particular it has been shown for the cases considered to what extent LMO's are transferable between structurally related molecules, if the total energy is taken as the measure of transferability. Additional approximations have been suggested: a) the truncation of the basis set for the expansion of the MO's in the region of interaction, b) the similar approximation for the projection operator for orthogonality, c) the neglect of the nonorthogonality of the MO's in the calculation of the energy expectation value, and d) the point charge approximation for the calculation of the interaction energy between the two sets of transferred LMO's. These additional approximations can be justified too.

Because these approximations acquire validity for distances between the atomic centers involved of the order of two normal bond lengths, the sets  $\Gamma$  and  $\Delta$  must be chosen with care. Except for the molecule  $C_2H_6$  the approximation of truncating the basis set has been examined for the case of the basis functions which have large exponential parameters and are mainly used to construct the inner shell MO's. In this case the distance between the centers must be about 6 a.u. for the Li and Be atoms in order to obtain a parallel shift of the potential curves relative to the exact SCF result to within  $10^{-3}$  to  $10^{-4}$  a.u. or 0.6 to 0.06 kcal/mole. (It has been stated at which distances agreement with the SCF result is obtained; in many practical questions, however, such stringent conditions might be unnecessary.) The same approximation can certainly be made for any type of basis function and any type of LMO, only the distance between the centers or centroids, respectively, must be greater to obtain equally good results. It is possible that in larger molecules this situation might improve, because the MO's in the region of interaction can serve as a kind of buffer.

It has been found that the sets  $\Gamma$  and  $\Delta$  are best chosen to be identical. It proves to be disadvantageous to allow in the set  $\Delta$  degrees of freedom which are not allowed in the set  $\Gamma$ . In other words the projection operators  $O_\Gamma$  and  $O_\Delta$  should project into the same subspace. This result could have been anticipated from the bounds in Eq. (37). The operator determining the MO's in the region of interaction is consequently given by Eq. (46).

Two different sources exist for the nonorthogonality of the MO's. The transferred LMO's of the two fragments are nonorthogonal because they result from calculations on different molecules and because their expansion might have been truncated due to the deletion of some basis functions in each of the two fragments. The MO's in the region of interaction are nonorthogonal to the transferred LMO's due to the projection effected by the operator  $O$ . The neglect of the non-orthogonality in the calculation of the total energy seems to be valid for the same distances between the atomic centers as the truncation of the basis set becomes valid. That this approximation works reliably, i.e. introduces essentially

only a parallel shift of the potential curves is vital for the aim of saving computational time.

The point charge approximation for the calculation of the interaction energy between the two sets of transferred LMO's in the two fragments can only be valid for large distances where quantum mechanical effects of bonding play no role. It seems to be valid for distances equal to or larger than the ones necessary for the validity of the other approximations discussed above. It should be mentioned that this approximation can easily and in various ways be refined. A rather crude but extremely simple version has been examined in this work, which allows to reduce the computational expense to a " $N^2$  law", where  $N$  is the number of basis functions, for a significant part of the calculation leaving the " $N^4$  law" only in the region of interaction. The present version of the approximation has given reasonable results in most cases considered, but its reliability in more complicated cases still remains to be demonstrated [51].

The theory of molecules in molecules discussed in the present article appears to be a promising starting point for reliable and time saving computations on larger molecules, but more experience is necessary with the approximations involved. This will supply the information where the method has to be ameliorated.

*Acknowledgement.* The calculations were performed on an IBM 360/91 computer while the author was at the Max-Planck-Institut für Physik und Astrophysik, München. The author would like to thank Professor Dr. L. Biermann and Dr. G. Dierksen for the opportunity of working in the institute and for using its facilities. Financial support of the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

## References

1. Gombas, P.: Pseudopotentiale. Wien, New York: Springer-Verlag 1967
2. Phillips, J.C., Kleinman, L.: Phys. Rev. **116**, 287 (1959)
3. Austin, B.J., Heine, V., Sham, L.J.: Phys. Rev. **127**, 276 (1962)
4. Weeks, J.D., Rice, S.A.: J. Chem. Phys. **49**, 2741 (1968); Weeks, J.D., Hazi, A., Rice, S.A.: Adv. Chem. Phys. **15**, 283 (1969)
5. Arai, T.: Rev. Mod. Phys. **32**, 370 (1960), and references contained in this review
6. Ellison, F.O.: J. Am. Chem. Soc. **85**, 3540 (1963) and several subsequent papers
7. Adams, W.H.: J. Chem. Phys. **34**, 89 (1961); **37**, 2009 (1962); **42**, 4030 (1965)
8. Adams, W.H.: Chem. Phys. Letters **11**, 441 (1971)
9. Adams, W.H.: Chem. Phys. Letters **11**, 71 (1971); applications to be published
10. Ohno, K., Inokuchi, H.: Theoret. chim. Acta (Berl.) **26**, 331 (1972)
11. McWeeny, R.: Proc. Roy. Soc. (London) A **223**, 306 (1954)
12. Lykos, P.G., Parr, R.G.: J. Chem. Phys. **24**, 1166 (1956)
13. Huzinaga, S., Cantu, A.A.: J. Chem. Phys. **55**, 5543 (1971)
14. von Niessen, W.: J. Chem. Phys. **55**, 1948 (1971)
15. Gilbert, T.L.: "Molecular orbitals in chemistry, physics, and biology", ed. by Löwdin, P.O., Pullman, B., p. 405. New York: Academic (1964)
16. Schlosser, H.: J. Chem. Phys. **53**, 4035 (1970); **55**, 5453, 5459 (1971); Schlosser, H., Gilbert, T.L.: J. Chem. Phys. **56**, 8 (1972)
17. Gilbert, T.L.: Phys. Rev. A **6**, 580 (1972)
18. Schlosser, H.: J. Chem. Phys. **57**, 4332, 4342 (1972)
19. Anderson, P.W.: Phys. Rev. Letters **21**, 13 (1968); Phys. Rev. **181**, 25 (1969)
20. Peters, D.: J. Chem. Phys. **46**, 4427 (1967)
21. Peters, D.: J. Chem. Phys. **51**, 1559, 1566 (1969)

22. Lennard-Jones, J.: Proc. Roy. Soc. (London) A **198**, 1, 14 (1949); Hall, G. G., Lennard-Jones, J.: *ibid.* A **202**, 155 (1950); Lennard-Jones, J., Pople, J. A.: *ibid.* A **202**, 166 (1950); Hall, G. G.: *ibid.* A **202**, 336 (1950); Pople, J. A.: *ibid.* A **202**, 323 (1950); Hall, G. G., Lennard-Jones, J.: *ibid.* A **205**, 357 (1951); Hall, G. G.: *ibid.* A **205**, 541 (1951); Hurley, A. C., Lennard-Jones, J., Pople, J. A.: *ibid.* A **220**, 446 (1953)
23. Boys, S. F.: Rev. Mod. Phys. **32**, 296 (1960); Foster, J. M., Boys, S. F.: *ibid.* **32**, 300 (1960); Boys, S. F.: "Quantum theory of atoms, molecules, and the solid state", ed. by Löwdin, P. O., p. 253. New York: Academic 1966
24. Edmiston, C., Ruedenberg, K.: Rev. Mod. Phys. **35**, 457 (1963); J. Chem. Phys. **43**, S 97 (1967); Ruedenberg, K.: "Modern quantum chemistry", ed. by Sinanoglu, O., Vol. I, p. 85. New York: Academic 1965
25. England, W., Salmon, L. S., Ruedenberg, K.: "Topics in current chemistry", Vol. 23, p. 31. Berlin: Springer 1971
26. Magnasco, V., Perico, A.: J. Chem. Phys. **47**, 971 (1967); **48**, 800 (1968)
27. Weinstein, H., Pauncz, R., Cohen, M.: Advances in mol. phys., Vol. 7, p. 97. New York: Academic 1971
28. Peters, D.: J. Chem. Soc. **2003**, 2105, 4017, 1963; **2901**, 2908, 2916, 1964; **3026**, 1965; **644**, 652, 656, 1966
29. von Niessen, W.: J. Chem. Phys. **56**, 4290 (1972); Theoret. chim. Acta (Berl.) **27**, 9 (1972); Theoret. chim. Acta (Berl.) **29**, 29 (1973)
30. Rothenberg, S.: J. Chem. Phys. **51**, 3389 (1969); J. Am. Chem. Soc. **93**, 68 (1970)
31. Switkes, I., Lipscomb, W. N., Newton, M. D.: J. Am. Chem. Soc. **92**, 3837, 3847 (1970)
32. Newton, M. D., Switkes, I.: J. Chem. Phys. **54**, 3179 (1971)
33. Levy, M., Stevens, W. J., Shull, H., Hagstrom, S.: J. Chem. Phys. **52**, 5483 (1970)
34. Adams, W. H.: J. Am. Chem. Soc. **92**, 2198 (1969); J. Chem. Phys. **57**, 2340 (1972)
35. Peters, D.: J. Am. Chem. Soc. **94**, 707 (1972)
36. England, W., Gordon, M. S.: J. Am. Chem. Soc. **93**, 4649 (1971); **94**, 4818, 5168 (1972); Chem. Phys. Letters **15**, 59 (1972)
37. Epstein, I. R.: J. Chem. Phys. **53**, 4425 (1970); Eisenberger, P., Marra, W. C.: Phys. Rev. Letters **27**, 1413 (1971); Roux, M., Epstein, I. R.: Chem. Phys. Letters **18**, 18 (1973)
38. Newton, M. D., Boer, F. P., Lipscomb, W. N.: J. Am. Chem. Soc. **88**, 2353, 2361, 2367, 2384 (1966); Boyd, D. B., Lipscomb, W. N.: J. Chem. Phys. **48**, 4955, 4968 (1968); Tossell, J. A., Lipscomb, W. N.: J. Am. Chem. Soc. **94**, 1505 (1972)
39. Bader, R. F. W., Beddall, P. M.: Chem. Phys. Letters **8**, 29 (1971); Bader, R. F. W., Beddall, P. M., Cade, P. E.: J. Am. Chem. Soc. **93**, 3095 (1971)
40. Christoffersen, R. F., Shipman, L. L., Maggiora, G. M.: Intern. J. Quantum Chem. **5**, 143 (1971)
41. Nelander, B.: J. Chem. Phys. **55**, 2949 (1971)
42. McWeeny, R., Del Re, G.: Theoret. chim. Acta (Berl.) **10**, 13 (1968)
43. Letcher, J. H., Dunning, T. H.: J. Chem. Phys. **48**, 4538 (1968); Urland, M. L., Dunning, T. H., Van Wazer, J. R.: *ibid.* **50**, 3208 (1969); Sovers, O. J., Kern, C. W., Pitzer, R. M., Kurplus, M.: *ibid.* **49**, 2592 (1968); Petke, J. D., Whitten, J. L.: *ibid.* **51**, 3166 (1969); Hoyland, J. R.: J. Am. Chem. Soc. **90**, 2227 (1968); J. Chem. Phys. **50**, 473 (1969); Letcher, J. H.: J. Chem. Phys. **54**, 3215 (1971)
44. Weinstein, H., Pauncz, R.: Chem. Phys. Letters **14**, 161 (1972)
45. Diner, S., Malrieu, J. P., Claverie, P.: Theoret. chim. Acta (Berl.) **13**, 1, 18 (1969); Diner, S., Malrieu, J. P., Jordan, F., Gilbert, M.: *ibid.* **15**, 100 (1969); Jordan, F., Gilbert, M., Malrieu, J. P., Pincelli, U.: *ibid.* **15**, 211 (1969)
46. Roothaan, C. C. J.: Rev. Mod. Phys. **23**, 69 (1951)
47. Löwdin, P. O.: Phys. Rev. **139**, A 357 (1965)
48. Roothaan, C. C. J.: Rev. Mod. Phys. **32**, 179 (1960)
49. Huzinaga, S.: J. Chem. Phys. **51**, 3971 (1969)
50. Löwdin, P. O.: Phys. Rev. **97**, 1474, 1490 (1955)
51. von Niessen, W.: in preparation
52. Hylleraas, E. A., Undheim, B.: Z. Physik **65**, 759 (1930); McDonald, J. K. L.: Phys. Rev. **43**, 830 (1933)
53. Huzinaga, S.: J. Chem. Phys. **42**, 1293 (1965)
54. Clementi, E.: IBM J. Res. Develop. Suppl. **9**, 2 (1965)



55. Löwdin, P. O.: J. Chem. Phys. **18**, 365 (1950)
56. Ransil, B. J., Sinai, I. J.: J. Chem. Phys. **46**, 4050 (1967)
57. Clementi, E.: J. Chem. Phys. **46**, 4737 (1967)
58. Veillard, A.: Chem. Phys. Letters **3**, 128, 565 (1969)
59. Weiss, S., Leroi, G. E.: J. Chem. Phys. **48**, 962 (1968)
60. Sovers, O. J., Kern, C. W., Pitzer, R. M., Karplus, M.: J. Chem. Phys. **49**, 2592 (1968)

Dr. W. von Niessen  
Lehrstuhl für Theoretische Chemie  
Technische Universität München  
D-8000 München 2, Arcisstr. 21  
Federal Republic of Germany



## SCF Dirac-Hartree-Fock Calculations in the Periodic System

### II. Binding Energies and First Ionization Potentials for *s*, *p*, and *d* Elements from $Z = 1$ to $Z = 120$

Jaromír Malý and Michel Hussonnois

Institut de Physique Nucléaire, Division de Radiochimie, Université de Paris XI, Centre d'Orsay,  
9t 406-Orsay, France

Received February 16, 1973

Binding energies calculated by DHF method were compared with modified DFS method calculations and experimental values. First ionization potentials of all elements from  $Z = 1$  to  $Z = 120$  (excluding the lanthanide and actinide series) were obtained from DHF values. These calculated values were compared with spectroscopically determined first ionization potentials for the region  $Z = 1$  to  $Z = 88$ . The obtained ratios of DHF calculated and experimental values in the  $Z \leq 88$  region (correlation ratios) were extrapolated for 104–120 elements and used in correcting calculated DHF eigenvalues to obtain expected values for the first ionization potential in this region.

*Key words:* Ionization potentials of atoms

#### 1. Introduction

Calculations have recently been published to the first ionization potentials of 104–120 element and 156–172 element [1, 2] (together with data for Ir, Os, Au and Hg), derived from the values of Dirac-Fock-Slater (DFS) calculations.

In our previous work we published calculations of all elements from  $Z = 1$  to  $Z = 120$ , using the more complex Dirac-Hartree-Fock method (DHF) [3].

The accuracy of the DHF method is generally greater than that of the DFS method used in [1] and [2], therefore, the comparison of the first ionization potentials from both methods is valuable. Having eigenvalues for all elements from H to 120 element from our DHF calculations, we also performed a number of additional calculations in the known regions, checking the accuracy of our calculation of binding energies from different subshells by two usual methods: as eigenvalues (method A) and as the difference between total energies of the  $1^+$  ion and atom (method B). Comparison of our values with those from modified DFS method [4] is also of interest.

This work, as part of our study of the periodic system by the DHF method, is based on the idea that any valuable extrapolation in unknown parts of the periodic system (beyond  $Z > 103$ ) must be supported by extensive calculations in the known region, using the same computer program.

## 2. Eigenvalues and Binding Energies

The general DHF method, as derived for many electron atoms [5, 6, 8] solves a set of integro-differential DHF equations including electric, magnetic and retardation interaction of each electron with all electrons. In this method the ionization potential (i.e. the electrons binding energy on each subshell with a given set of quantum numbers  $(n, l, j)$ ), is directly equal to the eigenvalue  $\epsilon_{nlj} = B_{nlj}$ . This supposes that removing an electron from the  $(n, l, j)$  subshell does not affect the other subshells (method of "frozen orbitals" according to Koopman's theorem [7]). This method we will refer to as "method A". The other method is that of calculating binding energy  $B_{nlj}$  as the difference between the total energy of the atom and corresponding  $1^+$  ion. This method, which we will refer to as "method B", includes rearranging effects on energy levels of all electrons in the atom after its ionization and should generally give results closer to those experimentally obtained. In our calculations we used this general schema derived in [6, 13 and 8] in formulae. We omit magnetic and retardation terms and used formulae and program as described previously in [3]. Omitting magnetic and retardation terms is not important for all outer shells in the atom - if these terms are included, they change the eigenvalues less than 1% (as shown in the case of Hg [8]). For "method B" we have from [3] -TE (average total energy - "Hartree Type") or -AE (average energy - "Slater type") available, both defined previously in [3].

With the DFS method a similar set of integrodifferential Dirac-Hartree-Fock equations is solved as in [3] (see Eqs. (21-25) in [3]) but all terms representing different potentials from direct and exchange interaction of electrons (bound on different subshells) and containing  $Y^E$ ,  $Y_c^v$ ,  $W_Q$  and  $W_P$  in [3], are replaced by potential  $V(r)$ :

$$V(r) = -\frac{Z}{r} + \frac{1}{r} \int_0^r \varrho(s) ds + \int_r^\infty \frac{\varrho(s)}{s} ds + V_{ex}(r). \quad (1)$$

This potential term is the same for all electrons and contains radial electron density  $\varrho$ , the square value of the radial part of the Dirac wave function:

$$\varrho(r) = \sum_j [P_j^2(r) + Q_j^2(r)]. \quad (2)$$

In this DFS model the negatively charged electron density  $\varrho$  interacts with the positively charged nucleus containing  $Z$  protons. Electron density  $\varrho$  is also interacting electrically with itself as expressed by the exchange potential  $V_{ex}$ :

$$V_{ex}(r) = -(C/r) [81 r^n \varrho^m(r) / 32 \pi^2]^{\frac{1}{3}}. \quad (3)$$

Formula (3) was derived by Slater [9] with  $C=1$ ,  $n=m=1$ . Gaspar [10] derived formula (3) with  $C=2/3$ ,  $n=m=1$  in a somewhat different way and Rosen and Lindgren [4] derived the modified DFS method (MDFS) by using formula (3) as a parametrical expression, when values  $C$ ,  $n$  and  $m$  were obtained by variational DFS calculation, minimizing the total energy. In [4] the set of parameters  $C$ ,  $n$ ,  $m$ , was found giving minimal total energy [called parameters of optimized potential in (3)]. For heavy atoms the optimized potential  $V_{ex}$  is

given with parameters  $C \simeq 2/3$ ,  $n=m=1$ , as in [10]. Because of the potential approximation (1), eigenvalues obtained by solving DFS integrodifferential equations do not exactly obey Koopman's theorem [7] as mentioned in [4]. The correct binding energies of subshell  $i$ , when calculated by "method A", are given in MDFS calculations [4]:

$$B_i = -\epsilon_i - \delta\epsilon_i \quad (4)$$

where

$$\delta\epsilon_i = \sum_j \langle ij|g|ij \rangle - \langle i|V(r) + Z/r|i \rangle \quad (5)$$

is the term (see details in [4]) which corrects the inexact eigenvalue  $\epsilon_i$  to binding energy  $B_i$ . The first term (5) contains summations of Slater integrals  $F^k$ ,  $G^k$  with proper coefficients, similar to those described in Eq. (18) of the exact DHF method [3]. The difference is that here DFS wave functions (not exact) are used in place of exact DHF wave functions, when the Slater integrals  $F^k$ ,  $G^k$  or the second term in (5) are calculated. With correction (5) the MDFS method expresses total energy as [4]:

$$E_{\text{tot}} = \sum_i \epsilon_i - \frac{1}{2} \sum_i \langle i|V(r) + Z/r|i \rangle + \frac{1}{2} \sum_i \delta\epsilon_i. \quad (6)$$

The normal DFS method (non modified) calculates  $E_{\text{tot}}$  using (6), omitting the term  $\frac{1}{2} \sum_i \delta\epsilon_i$ , which yields higher  $E_{\text{tot}}$  values (with smaller absolute values) than the MDFS or DHF methods.

Generally, the MDFS method can be treated as some approximative DHF method: it calculates  $E_{\text{tot}}$  with (6), i.e. using correct sums of Slater integrals  $F^k$ ,  $G^k$ , but with wave functions self consistently calculated by DFS method using the same approximative potential  $V(r)$  as defined in (1, 2, 3) when yielding non exact eigenvalues. We will see further that this approximation is the most valuable from DFS methods giving results close to DHF values.

When binding energies are calculated by DHF, MDFS or DFS method the total energy result corresponds to the barycentrum of the calculated electron configuration. Spectroscopically, this means that the calculated total energy should be compared with the barycentrum position of all terms included in a given electron configuration. The barycentrum position above ground term,  $\Delta_{SP}$ , is calculated according to the prescription in [11], p. 322, by the formula (for terms denoted in  $LS$  coupling in non-relativistic classification):

$$\Delta_{SP} = \frac{\sum_{L,S} \sum_{J_{L,S}} (2J_{L,S} + 1) E(J_{L,S})}{\sum_{L,S} \sum_{J_{L,S}} (2J_{L,S} + 1)}. \quad (7)$$

This formula means we take the energy  $E(J_{L,S})$  of each term ( $^{2S+1}L_J$ ) (with  $L$  denoted as  $S, P, D, F, \dots$  symbol) with the statistical weight  $(2J_{L,S} + 1)$  and sum all such contributions over all terms of the configuration. Here the factor  $(2J_{L,S} + 1)$  is equal to the number of Zeeman lines, in which can split the multiplet

$25+1L$  having quantum number  $J_{L,S}$  (it is equal to the number of separate wave functions of this multiplet).

In an exact comparison of DHF or DFS binding energies with experimental values, the experimental ionization potential (the difference of ground terms of atom and ion) must be corrected by  $\Delta_{sp}$  of atom and ion.

For the inner subshells (corresponding to X-ray levels) the spin orbit interaction prevails and their binding energies calculated by method "A" or "B" can be compared directly with X-ray levels.

### 3. Results for Binding Energies

Results for ionization potential of several  $1^+$  ions of Na, Mg, Al and systematically all  $1^+$  ions of Th are presented in Table 1 and 2. In these tables total energies as  $-TE$  or  $-AE$  values were first calculated by the DHF method, as described in [3], each in the electronic configuration of ground state atom and  $1^+$  ion (as shown in columns 1 and 2 of Tables 1 and 2). From the calculated values of  $-TE$  and  $-AE$  ("Hartree type" and "Slater type" total energies respectively) ionization potentials are calculated as DHF eigenvalues (marked "method A" in tables) and as the difference of  $-TE$  or  $-AE$  values of the atom and corresponding  $1^+$  ion (marked "method B"). These values are compared with experimental X-ray levels in the last column. As experimental levels, values from [12] are used, corrected for work function (+3.3 eV for Th and standardly 4.4 eV for all other values, corresponding to the work function of Cu slit, see [12]).

Table 1. Ionization potentials of  $1^+$  ions for Na, Mg, Al

Atom or ion	Electron configuration	Total energy DHF calculated in AU, -TE or -AE values	$1^+$ ionized subshell	Ionization potential in AU from -TE or -AE values (method B)	Eigenvalues (AU) (method A)	Experimental X-ray level [12] AU
Na	(Ne) $3s^1 + = (Na^0)$	162.0783				
Na $^{1+}$	(Na $^0$ , $1s^1 +$ )	122.4329	$1s +$	39.6454 (1.0016)	40.54489 (1.0243)	39.5827
Na $^{1+}$	(Na $^0$ , $2s^1 +$ )	159.4309	$2s +$	2.6474 (1.064)	2.80541 (1.127)	2.4893
Na $^{1+}$	(Na $^0$ , $2p^1 -$ )	160.7286	$2p -$	1.3497 (1.034)	1.52204 (1.166)	1.3053
Na $^{1+}$	(Na $^0$ , $3s^1 +$ )	161.8961	$3s +$	0.1822 (0.964)	0.18234 (0.965)	0.1890*
Mg	(Ne) $3s^2 + = (Mg^0)$	199.9353				
Mg $^{1+}$	(Mg $^0$ , $1s^1 +$ )	151.7404	$1s +$	48.1949 (1.0010)	49.12654 (1.0204)	48.1464
Mg $^{1+}$	(Mg $^0$ , $2s^1 +$ )	196.3418	$2s +$	3.5935 (1.042)	3.78017 (1.096)	3.4490
Mg $^{1+}$	(Mg $^0$ , $2p^1 -$ )	197.8544	$2p -$	2.0809 (1.0124)	2.28833 (1.113)	2.0554
Al	(Ne) $3s^2 + 3p^1 - = (Al^0)$	242.3315				
Al $^{1+}$	(Al $^0$ , $1s^1 +$ )	184.6739	$1s +$	57.6576 (1.0026)	58.63307 (1.0196)	57.5080
Al $^{1+}$	(Al $^0$ , $2s^1 +$ )	237.6271	$2s +$	4.7044 (1.048)	4.92893 (1.098)	4.4896

\* Optical data from [15].



The ratio of DHF calculated values to experimental X-ray levels is indicated in parentheses, showing the exactness of calculations. From these tables it is visible that ionization potential calculated by "method B" agree better with the experiment than potentials calculated by "method A". "Method B" values for  $1s+$  shell in the cases Na, Mg, Al are most exact (no relativistic effect is visible on this shell for lower  $Z$ ), differing only 0.1–0.26% from the experiment, with all other values less exact, but still differing only 3–6% from the experiment. In the case of Th by "method B" error on  $1s+$  and  $2s+$  values is 0.7% (due to the point nucleus approximation) and drops to 0.23% on the  $2p+$  values. On all levels with main quantum numbers 3 and 4, agreement is very good for levels 3 (giving errors from 0.71–0.16%) and good for levels 4 (within  $\sim 1\%$ ). Agreement is not very good for levels 5, 6 and 7. "Method A" is systematically giving somewhat higher results than "method B". However for the last shells ( $3s+$  in Na or 6th and 7th in Th) our results from "A" are closer to experiments than from "B".

In the case of Th one can observe systematics in the deviations between experiments and "method B" results. Shells with 2 electrons ( $s+$  or  $p-$ ) show the greatest deviation. As the number of electrons on closed shells increases, the difference decreases from  $3s+$  to  $3d+$  or from  $4s+$  to  $4f+$ .

Table 3. Comparison of calculated values for  $\text{Hg}^{80}(\text{AU})$ 

Ionized subshell	Binding energies			Correc- tion [8] $A_{\text{C}, \text{R}}$	Binding energies	
	Eigenvalues DHF This work (method A)	MDFS [4] (method A)	MDFS [4] (method B)		Expected values (MDFS – $A_{\text{C}, \text{R}}$ )	Experimental values [12]
$1s+$	3076.1399	3076.15	3072.70 (1.0061)	– 11.374	3061.33 (1.00234)	$3054.19 \pm 0.030$
$2s+$	550.5322	550.50	548.41 (1.0053)	– 1.548	546.86 (1.00247)	$545.51 \pm 0.035$
$2p$	526.8543	526.85	524.55 (1.0042)	– 2.407	522.14 (0.99962)	$522.34 \pm 0.025$
$2p+$	455.1315	455.13	453.04 (1.0042)	– 1.542	451.50 (0.99978)	$451.60 \pm 0.015$
$3s+$	133.1796	133.14	131.90 (1.0064)	– 0.358	131.54 (1.00366)	$131.06 \pm 0.040$
$3p-$	122.6415	122.61	121.32 (1.0056)	– 0.503	120.82 (1.00141)	$120.65 \pm 0.045$
$3p+$	106.5334	106.52		– 0.336		$104.80 \pm 0.015$
$3d-$	89.4262	89.41	88.15 (1.0039)	– 0.306	87.84 (1.00034)	$87.81 \pm 0.010$
$3d+$	86.0061	86.00		– 0.239		$84.50 \pm 0.010$
$4s+$	30.6701	30.66	30.04 (1.0156)	– 0.093	29.95 (1.0125)	$29.58 \pm 0.035$
$4p-$	26.1297	26.12	25.50 (1.0185)	– 0.130	25.37 (1.0132)	$25.04 \pm 0.085$
$4p+$	22.1865	22.19		– 0.080		$21.15 \pm 0.050$
$4d-$	14.7954	14.80	14.27 (1.0142)	– 0.064	14.21 (1.0099)	$14.07 \pm 0.035$
$4d+$	14.0473	14.05		– 0.049		$13.38 \pm 0.045$
$4f-$	4.4644	4.476	3.999 (1.0176)	– 0.028	3.971 (1.0104)	$3.93 \pm 0.020$
$4f+$	4.3025	4.315		– 0.022		$3.79 \pm 0.020$
$5s+$	5.1088	5.125	4.896 (1.0667)	– 0.018	4.878 (1.0627)	$4.59 \pm 0.045$
$5p-$	3.5402	3.553	3.344 (1.0718)	– 0.022	3.322 (1.0647)	$3.12 \pm 0.045$
$5p+$	2.8418	2.856	2.687 (1.1785)	– 0.013	2.674 (1.1728)	$2.28 \pm 0.045$
$5d-$	0.65009	0.659	0.543 (0.884)	– 0.006	0.537 (0.875)	$0.614 \pm 0.050^*$
$5d+$	0.57382	0.584	0.483 (0.886)	– 0.004	0.479 (0.879)	$0.545 \pm 0.050^*$
$6s+$	0.32863	0.340	0.312 (0.813)	– 0.002	0.310 (0.813)	$0.384 \pm 0.050^*$

\* Data from optical system of [14].



The explanation for this behaviour could be connected with the fact that magnetic and retardation terms are omitted in formulae [3].

We may demonstrate this in the case of  $\text{Hg}^{80}$ , calculated exactly in [8] ("method A") with all magnetic and retardation terms included. In Table 3 we compared our results by "method A" for  $\text{Hg}^{80}$  (column 1) with MDFS corrected binding energies obtained by methods "A" and "B" from paper [4] with optimized potential – using formula (5) for "A" and the difference of atom and ion total energies from "B", as defined in formula (6).

Our results from "method A" and [4] show excellent agreement, proving that MDFS results of (4) are essentially very close to our DHF results. The ratio of "method B" results in Table 3 to experimental values, as presented in parentheses, show similarity to equivalent ratios in the case of our Th + calculations (see Table 2, "method B"). We also compared our DHF values – eigenvalues for Cu, Kr, J, Eu and U from [3] with binding energies of "method A" calculated in [4] using Eqs. (1–6). The agreement with our DHF eigenvalues was also very good in all levels of these cases (from  $Z = 29$  to  $Z = 92$ ), as in the case of  $\text{Hg}^{80}$  in Table 3. From this we can confirm the conclusion drawn in [4] that the MDFS method gives (up to  $Z = 92$ ) practically the same results as DHF. Furthermore, if we use (in Table 3) the sum of magnetic (Gaunt) and retardation terms calculated exactly for  $\text{Hg}^{80}$  in [8] as correction  $\Delta_{G+R}$  (Table 3, column 6), which lowers the binding energies "B", we may obtain the expected experimental values (column 6). When we compare the expected values with experimental values (its ratio in parentheses, sixth column) we see agreement is greatly

Table 4. Ionization potential of polyvalent ions

Ionization process	Electron structure of ion in ground state (DHF)	Calculated ionization potential (DHF) in AU			Experimental ionization potential (AU) [15]	
		From -TE difference	From -AE difference	From eigenvalues	Directly measured	Barycentrum difference
Na $\rightarrow$ Na <sup>1+</sup>	(Ne)	0.1822	0.1822 (0.964)	0.18234 (0.965)	0.1890	0.1890 (0.965)*
Mg $\rightarrow$ Mg <sup>2+</sup>	(Ne)	0.7850	0.7850 (0.941)		0.8340	0.8340
Al $\rightarrow$ Al <sup>3+</sup>	(Ne)	1.8794	1.8794 (0.960)		1.9584	1.9581
Si $\rightarrow$ Si <sup>1+</sup>	(Ne)	0.2486	0.2486 (0.829)	0.26695 (0.891)	0.2997	0.2859 (0.934)
	$3s^2 + 3p^1 -$					
Si $\rightarrow$ Si <sup>2+</sup>	(Ne) $3s^2 +$	0.8263	0.8263 (0.917)	0.85339 (0.947)	0.9007	0.8860 (0.963)
Si $\rightarrow$ Si <sup>3+</sup>	(Ne) $3s^1 +$	2.0032	2.0032 (0.939)	2.03784 (0.956)	2.1322	2.1175 (0.962)
Si $\rightarrow$ Si <sup>4+</sup>	(Ne)	3.6472	3.6472 (0.962)	3.68269 (0.971)	3.7920	3.7773 (0.975)
Th $\rightarrow$ Th <sup>1+</sup>	(Rn)	0.195	0.194 (0.758)	0.20651 (0.807)	0.256 [16]	
	$6d^2 - 7s^1 +$				0.276 [17]	
Th $\rightarrow$ Th <sup>2+</sup>	(Rn) $6d^2 -$	0.602	0.601 (0.860)	0.61602 (0.881)	0.699 [15]	
Th $\rightarrow$ Th <sup>3+</sup>	(Rn) $5f^1 -$	1.253	1.251 (0.872)	1.27782 (0.891)	1.434 [15]	
Th $\rightarrow$ Th <sup>4+</sup>	(Rn)	2.232	2.237 (0.897)	2.31642 (0.929)	2.493 [15]	

\* In the last column are ratio:  $\frac{\text{IP from eigenvalues}}{\text{IP from barycentrum difference}}$ ; other values are compared to IP directly measured.

improved. With the exclusion of ionization levels  $1s+$  and  $2s+$  (deviation caused by point nucleus approximation and Lamb shift) the expected binding energies agree almost exactly with the experiment for the main quantum levels 2, 3 and 4. However, even after deducing  $\Delta_{G+R}$ , one can see that for quantum levels 5 and 6, agreement is not very good with data from X-ray or optical spectra. This is equally valid for more exact DHF calculations of Hg eigenvalues [8] which includes the magnetic and retardation terms and finite nucleus.

The remaining difference (after deducing  $\Delta_{G+R}$ ) is, perhaps, caused by correlation effects and by approximations used in Dirac-Hartree-Fock Eqs. in [3] or [8] - e.g. by using the nonexact  $\Gamma_{ikj}$  coefficient for open shells, which is valid exactly for an atom with only all closed shells. It is apparent that calculation of the  $1^+$  ionized shell (by "method B") from a subshell with many electrons (as  $nd+$  or  $nf+$  shell) is more exact than calculation of the  $1^+$  ion from  $s+$  or  $p-$  subshells (containing only 2 electrons when the shell is full) - when the same  $\Gamma_{ikj}$  coefficients valid for closed shells are used (see inner subshells in Table 3 with  $\Delta_{G+R}$  correction, or in Table 2 without such correction).

In Table 4, similar calculations for  $1^+$ ,  $2^+$ ,  $3^+$ ,  $4^+$  ions are shown as in Tables 1, 2, 3 for  $1^+$  ions.

In DHF calculations it is visible that in all cases the accord is better in the region of valence electrons with eigenvalues, obtained from "method A", than with values obtained by "method B".

This accord can be somewhat improved when DHF results are compared with ionization potentials obtained from the barycentrum difference using calculations  $A_{SP}$  from formula (7) for atom and ion (see Table 4, last column).

Generally, Tables 1, 2, 3, 4 show the accuracy with which ionization potentials from DHF values can be calculated, using approximations as described in [3]. It is visible that accuracy of 0.1–2% can be achieved in comparison to X-ray levels, for all inner shells when the deviation of eigenvalues from experimental ionization potentials for valence electrons is in the range 2–10%.

#### 4. Results for the First Ionization Potentials

The preceding tables have shown that the first ionization potentials for valence electrons closest to experimental values can be found most simply from DHF eigenvalues with an accuracy of  $\sim 10\%$ . Therefore we use our eigenvalues from data calculated in [3] to compare first ionization potentials of all atoms in the periodic system with their experimental values (in  $Z=1-88$  region, see Table 5). In a previous paper [3], we found some discrepancies between DHF calculated and measured ground state electron configurations in cases Cr, Cu, Nb, Te and Pd. However, these discrepancies could be only apparent and need to be verified by the exact calculations of barycentrum positions of both concurrent electron states, according to (7), including the statistically weighted participation of (+) and (-) states in each open  $n, l$  subshell.

At present, therefore, we compare both possible configurations in Table 5. The ratios of the first ionization potential calculated by DHF to the experimental

Table 5. First ionization potential in the periodic system

Element	Z	Ionized level	Ionization potential (eV)			I.P. from barycentrum difference	Barycentrum correlation ratio
			Calculated DHF I.P.	Experimental I.P. [15]	Correlation ratio		
H	1	1s <sup>1</sup> +	13.598	13.598	1.000	13.598	1.000
He	2	1s <sup>2</sup> +	24.966	24.587	1.015	24.587	1.015
Li	3	1s <sup>1</sup> +	5.340	5.392	0.990	5.392	0.990
		1s <sup>2</sup> +	67.392				
Be	4	2s <sup>2</sup> +	8.412	9.322	0.902	9.322	0.902
		1s <sup>2</sup> +	128.733				
B	5	2p <sup>1</sup> -	8.426	8.298	1.016	8.297	1.016
		2s <sup>2</sup> +	13.459				
C	6	2p <sup>2</sup> -	10.599	11.260	0.941	10.710	0.990
		2s <sup>2</sup> +	19.496				
N	7	2p <sup>2</sup> -	13.323	14.534	0.917	13.220	1.008
		2p <sup>1</sup> +	14.673				
O	8	2p <sup>2</sup> -	16.326	13.618	1.200	15.911	1.026
		2p <sup>2</sup> +	17.327				
F	9	2p <sup>2</sup> -	19.614	17.422	1.126	18.652	1.052
		2p <sup>3</sup> +	20.124				
Ne	10	2p <sup>4</sup> +	23.069	21.564	1.070	21.564	1.070
		2p <sup>2</sup> -	23.194				
Na	11	3s <sup>1</sup> +	4.959	5.139	0.965	5.139	0.965
		2p <sup>4</sup> +	41.191				
Mg	12	3s <sup>2</sup> +	6.893	7.646	0.902	7.646	0.902
		2p <sup>4</sup> +	61.917				
Al	13	3p <sup>1</sup> -	5.711	5.986	0.954	5.977	0.955
		3s <sup>2</sup> +	10.722				
Si	14	3p <sup>2</sup> -	7.260	8.151	0.891	7.777	0.934
		3s <sup>2</sup> +	14.891				
P	15	3p <sup>2</sup> -	9.225	10.486	0.880	9.654	0.956
		3p <sup>1</sup> +	10.016				
S	16	3p <sup>2</sup> -	11.366	10.360	1.097	11.615	0.979
		3p <sup>2</sup> +	11.907				
Cl	17	3p <sup>2</sup> -	13.688	12.967	1.056	13.668	1.001
		3p <sup>3</sup> +	13.895				
Ar	18	3p <sup>4</sup> +	15.986	15.759	1.014	15.759	1.014
		3p <sup>2</sup> -	16.194				
K	19	4s <sup>1</sup> +	4.026	4.341	0.927		
		3p <sup>4</sup> +	25.823				
Ca	20	4s <sup>2</sup> +	5.339	6.113	0.874		
		3p <sup>4</sup> +	36.272				
Sc	21	4s <sup>2</sup> +	5.740	6.54	0.878		
		3d <sup>1</sup> -	9.512				
Ti	22	4s <sup>2</sup> +	6.068	6.82	0.890		
		3d <sup>2</sup> -	11.177				
V	23	4s <sup>2</sup> +	6.363	6.74	0.945		
		3d <sup>3</sup> -	12.655				
Cr	24 <sup>a</sup>	4s <sup>2</sup> +	6.638	6.766	0.980		
		3d <sup>4</sup> -	14.023				
Cr	24 <sup>b</sup>	4s <sup>1</sup> +	5.708	6.766	0.844		
		3d <sup>1</sup> +	8.532				
Mn	25	4s <sup>2</sup> +	6.904	7.435	0.930		
		3d <sup>1</sup> +	14.904				
Fe	26	4s <sup>2</sup> +	7.159	7.870	0.910		
		3d <sup>2</sup> +	16.105				

Table 5 (continued)

Element	Z	Ionized level	Ionization potential (eV)		
			Calculated DHF I.P.	Experimental I.P. [15]	Correlation ratio
Co	27	$4s^2 +$	7.408	7.86	0.943
		$3d^1 +$	17.261		
Ni	28	$4s^2 +$	7.650	7.635	1.002
		$3d^4 +$	18.378		
Cu	29 <sup>a</sup>	$4s^2 +$	7.888	7.726	1.020
		$3d^5 +$	19.464		
Cu	29 <sup>b</sup>	$4s^1 +$	6.657	7.726	0.861
		$3d^6 +$	12.898		
Zn	30	$4s^2 +$	8.122	9.394	0.864
		$3d^0 +$	20.524		
Ga	31	$4p^1 -$	5.707	5.999	0.952
		$4s^2 +$	11.773		
Ge	32	$4p^2 -$	7.060	7.899	0.894
		$4s^2 +$	15.505		
As	33	$4p^2 -$	8.800	9.81	0.898
		$4p^1 +$	9.268		
Se	34	$4p^2$	10.654	9.752	1.093
		$4p^2 +$	10.789		
Br	35	$4p^1 +$	12.360	11.814	1.045
		$4p^2$	12.360		
Kr	36	$4p^4 +$	13.987	13.999	0.999
		$4p^1 -$	14.731		
Rb	37	$5s^1 +$	3.809	4.177	0.911
		$4p^4 +$	21.653		
Sr	38	$5s^2 +$	4.930	5.695	0.866
		$4p^4 +$	29.366		
Y	39	$5s^2 +$	5.457	6.38	0.855
		$4d^1$	6.599		
Zr	40	$5s^2 +$	5.830	6.84	0.853
		$4d^2 -$	8.230		
Nb	41 <sup>a</sup>	$5s^2 +$	6.135	6.88	0.891
		$4d^1$	9.752		
Nb	41 <sup>b</sup>	$5s^1 +$	5.538	6.88	0.805
		$4d^4 -$	7.165		
Mo	42 <sup>a</sup>	$5s^2 +$	6.401	7.099	0.902
		$4d^4$	11.229		
Mo	42 <sup>b</sup>	$5s^1 +$	5.727	7.099	0.808
		$4d^1 +$	7.954		
Tc	43 <sup>a</sup>	$5s^1 +$	5.895	7.28	0.811
		$4d^2 +$	9.087		
Tc	43 <sup>b</sup>	$5s^2 +$	6.648	7.28	0.912
		$4d^1 +$	12.213		
Ru	44	$5s^1 +$	6.049	7.37	0.821
		$4d^3 +$	10.216		
Rh	45	$5s^1 +$	6.191	7.46	0.831
		$4d^4 +$	11.347		
Pd	46 <sup>a</sup>	$5s^1 +$	6.325	8.34	0.758
		$5d^3 +$	12.486		
Pd	46 <sup>b</sup>	$4d^6 +$	8.694	8.34	1.041
		$4d^4 -$	9.266		
Ag	47	$5s^1 +$	6.452	7.576	0.852
		$4d^6 +$	13.633		

Table 5 (continued)

Element	Z	Ionized level	Ionization potential (eV)		
			Calculated DHF I.P.	Experimental I.P. [15]	Correlation ratio
Cd	48	$5s^2 +$	7.657	8.993	0.852
		$4d^6 +$	19.270		
In	49	$5p^1 -$	5.492	5.786	0.949
		$5s^2 +$	10.699		
Sn	50	$5p^2 -$	6.631	7.344	0.892
		$5s^2 +$	13.762		
Sb	51	$5p^1 +$	8.147	8.641	0.943
		$5p^2 -$	8.197		
Te	52	$5p^2 +$	9.392	9.009	1.043
		$5p^2 -$	9.843		
I	53	$5p^3 +$	10.660	10.451	0.981
		$5p^2 -$	11.576		
Xe	54	$5p^4 +$	11.958	12.130	0.985
		$5p^2 +$	13.403		
Cs	55	$6s^1 +$	3.488	3.894	0.897
		$5p^4 +$	17.928		
Ba	56	$6s^2 +$	4.439	5.212	0.852
		$5p^4 +$	23.728		
La	57	$6s^2 +$	4.867	5.577	0.873
		$5d^1 -$	6.724		
Hf	72	$6s^2 +$	6.440	7.0	0.920
		$5d^2 -$	6.953		
Ta	73	$6s^2 +$	6.807	7.89	0.863
		$5d^3 -$	8.270		
W	74	$6s^2 +$	7.134	7.98	0.894
		$5d^4 -$	9.536		
Re	75	$6s^2 +$	7.477	7.88	0.948
		$5d^1 +$	9.704		
Os	76	$6s^2 +$	7.794	8.7	0.896
		$5d^2 +$	10.885		
Ir	77	$6s^2 +$	8.095	9.1	0.890
		$5d^3 +$	12.061		
Pt	78 <sup>a</sup>	$6s^1 +$	7.722	9.0	0.859
		$5d^5 +$	10.642		
Pt	78 <sup>b</sup>	$6s^2 +$	8.384	9.0	0.932
		$5d^4 +$	13.238		
Au	79	$6s^1 +$	7.948	9.225	0.859
		$5d^6 +$	11.647		
Hg	80	$6s^2 +$	8.937	10.437	0.857
		$5d^6 +$	15.606		
Tl	81	$6p^1 -$	5.805	6.108	0.948
		$6s^2 +$	11.921		
Pb	82	$6p^2 -$	6.904	7.416	0.932
		$6s^2 +$	14.907		
Bi	83	$6p^1 +$	6.995	7.286	0.962
		$6p^2 -$	8.754		
Po	84	$6p^2 +$	8.142	8.42	0.968
		$6p^2 -$	10.656		
At	85	$6p^3 +$	9.286		(0.969) <sup>c</sup>
		$6p^2 -$	12.638		
Rn	86	$6p^4 +$	10.438	10.748	0.970
		$6p^2 -$	14.709		

Table 5 (continued)

Element	Z	Ionized level	Ionization potential (eV)		
			Calculated DHF I.P.	Experimental I.P. [15]	Correlation ratio
Fr	87	$7s^1 +$	3.614		(0.914) <sup>c</sup>
		$6p^4 +$	15.402		
Ra	88	$7s^2 +$	4.527	5.279	0.858
		$6p^4 +$	20.126		
Ac	89	$6d^1 -$	5.100		
		$7s^2 +$	5.139		

<sup>a</sup> <sup>b</sup> <sup>c</sup> corresponds to two possible ground states (see [3]).

<sup>d</sup> Interpolated values

one (called correlation ratio) are presented in column 5. For comparison we present the ionization potential for the  $1^+$  ion of the two lowest lying subshells in each atom. From this one can see how both subshells are energetically close. Some  $p^-$  and  $p^+$  subshells, when both are full or close to be full, are really energetically very close, as visible in Table 5.

The agreement of experimental and calculated values is good, usually within  $\sim 10\%$ . For the first two periods (up to Ar) we also calculated ionization potentials for the barycentrum position of spectral lines of atom and ion, according to formula (7) using data of [15]. The agreement is substantially improved in some cases (see "barycentrum correlation ratio" when the difference of the ion and atom barycentrum is used as the experimental ionization potential). However, this comparison is not completely valid, because we correlate DHF values which were calculated for a full  $2p^-$  or  $3p^-$  shell and a non-filled  $2p^+$  or  $3p^+$  shell - with spectroscopic barycentrum difference, which contains the statistically weighted spectroscopic terms from a non-relativistic  $2p$  or  $3p$  shell according to (7) ( $LS$  coupling). This correlation supposes that  $p^-$  and  $p^+$  shells have the same ionization potential which is only a crude approximation here. For correct correlation more DHF calculations are needed. For example DHF values calculated for  $(2p^1 - 2p^3 +)$ ,  $(2p^4 +)$ ,  $(2p^2 - 2p^2 +)$  states in the case of oxygen (calculated in  $jj$  coupling) should be statically weighted similarly as in (7) to obtain the average energy of the  $(2p^4)$  oxygen state, using more complex formulae, see e.g. [22].

It is interesting to observe that correlation ratios have a systematic "periodic behavior", moving from the beginning to the end of each period, as visible from Fig. 1. From this figure one can make a fairly reasonable extrapolation along the  $Z$  axis for each chemical group of the periodic system to the region of 104–120 elements. In Fig. 1 it is visible that all  $s$  and  $p$  elements are forming a curve in each period with a maximum on the 6th group (O, S, Se, Te, Po). Good extrapolation for all  $s$  and  $p$  elements can be made by linear extension of their correlation ratios between the last two members of each chemical group. By this way extrapolation ratios were obtained for 112–120 element. As another possible extrapolation method for element 116, linear interpolation between

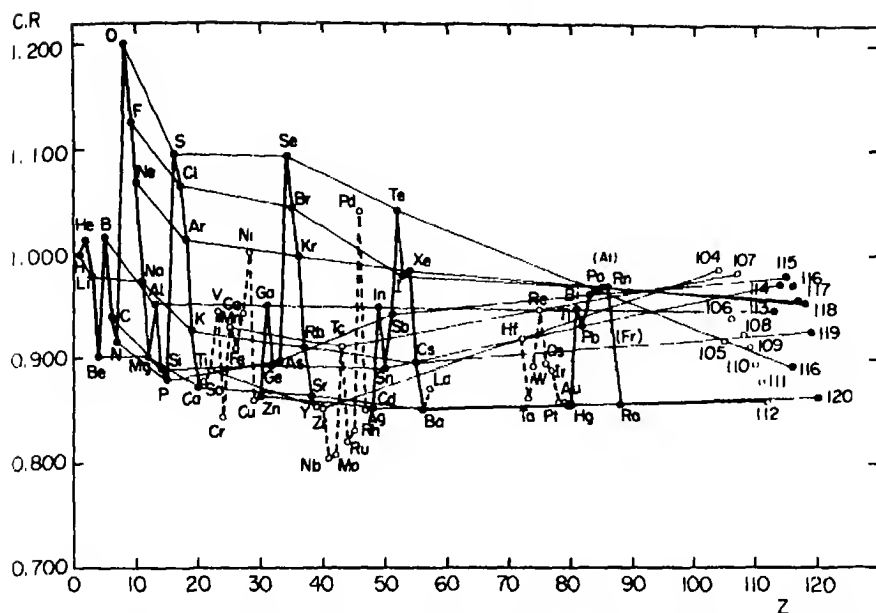


Fig. 1. Correlation ratios in the periodic system. ● *s* and *p* elements; ○ *d* elements. Values for 104, 107, 112–120 elements are linearly extrapolated correlation ratios. Values for 105, 106, 108 and 109 element are extrapolated with aid of Fig. 2. Points for 110, 111 element in Fig. 1 are interpolated between values for 109 and 112 element

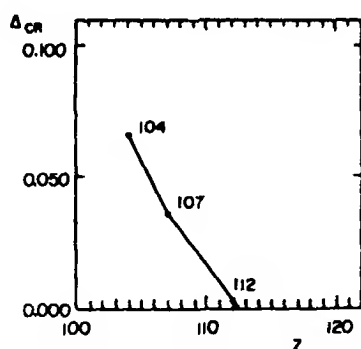


Fig. 2. Correction of correlation ratios for *d* elements  $\Delta_{C.R.}$  in the 104–112 element region

extrapolated values of 118 and 115 elements was used. To obtain values for elements 117 and 119, the interpolation between Ra and Rn and Po was used first, to obtain correlation ratios for At and Fr as a basis for extrapolation. For *d* elements direct extrapolation is possible for 104, 107 and 112 elements – they have systematically similar electron structure ( $xds^2$ ). For 105, 108, 109, 110 and 111 elements such extrapolation is impossible because their chemical analogues in the periodic system either do not have an electron structure analogical to them (Pt, Au) or do not have analogical electron structure just before and after the

Table 6. First ionization potential for  $Z = 104 - 120$  elements

Element $Z$	Ionized level	Ionization potential (eV)			Expected I.P.	
		Calculated DHF I.P.	Expected I.P.	Extrapolated correlation ratio	From DFS data [1]	Extrapolated from periodic system [20, 19]
104 <sup>a</sup>	$7p^1$	5.049	5.1			
	$6d^1 -$	8.070				
104 <sup>b</sup>	$6d^2 -$	5.770	5.9	0.986	5.1	
	$7s^2 +$	7.290				
105 <sup>a</sup>	$6d^1$	6.871	7.5	0.918	6.2	
	$7s^2 +$	7.874				
105 <sup>b</sup>	$7p^1$	5.518	6.0			
	$6d^2$	9.252				
106	$6d^3$	7.906	8.4	0.939	7.1	
	$7s^2 +$	8.408				
107	$6d^1 +$	7.289	7.4	0.983	6.5	
	$7s^2 +$	9.084				
108	$6d^2 +$	8.242	8.9	0.924	7.4	
	$7s^2 +$	9.751				
109	$6d^1 +$	9.173	10.1	0.912	8.2	
	$7s^2 +$	10.421				
110	$6d^3 +$	10.093	11.3	0.895	9.4	
	$7s^2 +$	11.106				
111	$6d^3 +$	11.005	12.5	0.880	10.3	
	$7s^2 +$	11.811				
112	$6d^6 +$	11.918	13.9	0.860	11.2	
	$7s^2 +$	12.549				
113	$7p^1$	7.464	7.9	0.946	7.5, 7.4 [21]	5.98 [20]
	$6d^6 +$	15.325				
114	$7p^2$	8.519	8.8	0.972	8.5, 8.5 [21]	
	$7s^2 +$	18.455				
115	$7p^1 +$	5.046	5.1	0.981	5.9	
	$7p^2$	11.491				
116	$7p^2 +$	6.170	6.6	0.933	6.8	
	$7p^2 -$	14.344				
117	$7p^3 +$	7.220	7.6	0.957	8.2	8.41 [20] 9.3 [19]
	$7p^2 -$	17.264				
118	$7p^3 +$	8.243	8.6	0.955	9.0	9.8 [19]
	$7p^2 -$	20.309				
119	$8s^1 +$	4.454	4.9	0.916	4.1	3.72 [20] 3.6 [19]
	$7p^3 +$	11.661				
120	$8s^2 +$	5.368	6.2	0.864	5.3	4.94 [20] 5.4 [19]
	$7p^3 +$	14.960				

lanthanide serie (Nb-Ta, Mo-W, Ru-Os, Rh-Ir, - giving only one point for extrapolation).

For the correlation ratios of  $d$  elements we suppose therefore that these can be obtained from correlation ratios of elements between Hf and Hg, by recalculating, according to Fig. 2. In this figure, the change of correlation ratios,  $\Delta_{CR}$ , between Hf and 104, Re and 107 and Hg and 112, obtained from direct extrapolation as shown in Fig. 1, is plotted. One can see in Fig. 2, that the



change  $\Delta_{CR}$  is approximately in linear dependence to  $Z$  between elements 104–112. Then for 105, 106, 108 and 109 elements we took the correlation ratio values for Ta, W, Os and Ir from Table 5 and corrected them by the corresponding  $\Delta_{CR}$  from Fig. 2. Such correlation ratios are also plotted in Fig. 1 (as isolated points). Data for 110 and 111 elements were then obtained as linear interpolation between 109 and 112 elements in Fig. 1.

The described extrapolation yields the "extrapolated correlation ratios" for the 104–120 element region, as shown in column 5 of Table 6. In this table DHF calculated ionization potentials (eigenvalues) from [3] are also presented for the last two subshells. For 116 element the mean value of both possible extrapolations is used (marked by two points in Fig. 1).

In column 4 of Table 6, the "expected I.P.", i.e. the expected experimental first ionization potential of each element, is presented, calculated using the extrapolated correlation ratio (or mean value of its two possible values) and DHF calculated I.P. as basic data. For comparison, in column 6, the expected I.P. from DFS calculations [1] are presented. In the last column some data [19] and [20] are presented, which were obtained by extrapolation from the trends in the periodic system (without DHF or DFS calculations).

## 5. Discussion

The comparison of our values sometimes shows substantial differences from DFS data [1] and even greater differences from simply extrapolated data [19, 20]. The differences between our values and data in [1] and [2] are caused by different approximations, used in solution of the Dirac equation. The eigenvalues and total energies in [1] and [2] are apparently obtained by the noncorrected DFS method. The theoretical descriptions of formulae in [1] do not present the corrections  $\delta\epsilon_i$ , as shown in Eqs. (4–6) – which could correct the DFS eigenvalues and DFS total energies to more exact values of the MDFS or DHF method. When we compare eigenvalues of valence electrons, published in [1] and [2] with ours, we find them very low because of this noncorrected DFS approximation.

Also the total energies of [1] are similarly noncorrected by formula (6) (as is apparent e.g. for the total energy of  $\text{Au}^{79}$  in [1] when compared with noncorrected DFS and corrected MDFS values for  $\text{Au}^{79}$  presented in [4], including the differences for point nucleus in [1] and finite nucleus in [4]). It is natural that the calculated ionization potentials of [1], obtained by "method B" from DFS noncorrected total energies of atom and  $1^+$  ion, generally will differ from MDFS or our DHF eigenvalues or ionization potentials (as shown in Table 5 and 6).

For correcting calculated ionization potentials to expected values, only a very briefly tested relation was used in [1]: DFS calculations of ionization potentials for Pt, Au and Hg were correlated to experimental values and gave very optimistic correlation values (found a difference 0.2 eV corresponding to our correlation ratio 0.98). On this basis, to calculate DFS ionization potentials in [1] and [2] + 0.2 eV was added in the  $d$  electron region and + 0.8 eV (estimated

similarly in the Pb region in [1]) in the  $p$  electron region, to correct the calculated ionization potentials to expected experimental values.

Our extrapolation in Figs. 1 and 2 shows, however, that a more complex correction procedure should be applied. In some cases this caused strong differences between our values and the values of ionization potentials calculated in [1]. If we compare the non-corrected ionization potentials as calculated in [1] (i.e. values from [1] after deduction of 0.2 eV or 0.8 eV correction respectively), we see relatively good agreement with our DHF eigenvalues in 115–120 element region. In the region 104–114 the DFS calculated ionization potentials (non-corrected) are systematically lower by  $\sim 1$  eV than our calculated DHF values. We believe that our values are more correct than in [1] and [2], and that they are, probably, generally accurate within  $\pm 5\%$  when correlation ratios can be extrapolated by the singly way. When two possible extrapolations are used, the accuracy is defined by both correlation ratios as limit values (probably within  $\pm 10\%$ , see element 116). Elements 104 and 105 are presented in two possible energetically close configurations, each of them using the same correlation ratios.

Generally from the data in Table 6, it is not possible to make a valuable prediction of the chemical behavior of unknown elements. However our results are valuable for mass spectroscopic behavior and some preliminary prediction could be done.

In mass spectrometrical separation of an ore in the search of a superheavy element in the 110–112 element region, one can see from Table 6 that a very strong (electron bombardment) ionization is necessary. The ionization potential of these elements is so high that they could easily be lost (nonionized) during mass separation, when similar Pt, At and Hg are well ionized in the  $1^+$  state and collected with good yield. The first ionization potential of 112 element is close to the ionization potential of Kr. From this fact, it is possible that element 112 and perhaps 111 would not follow well the chemistry of Au and Hg, but perhaps would behave mostly as noble metal. A more conclusive basis for prediction may only be given by complete DHF calculations of their different ions and based on the use of Born-Haber cycles.

On the other hand, the mass spectroscopic determination of 104 element should be much easier than for Hf, perhaps as in the cases of  $\text{Ca}^+$  and  $\text{Li}^+$ . The mass spectroscopic determination of 118 element should be as easy as determination of  $\text{Te}^+$  or  $\text{Po}^+$ .

The mass spectroscopic ionization and separation of element 119 and 120 should be much less easy than for Cs and Ba, probably similar to Na and Sr. Chemically, the 119 and 120 elements in solution could be probably much less basic than Cs and Ba, perhaps giving (120)  $\text{SO}_4$  soluble as  $\text{SrSO}_4$ .

*Acknowledgement* The authors wish to express their gratitude to Professor G. Bouissières for his encouragement and continuous support of this work.

## References

1. Fricke, B., Greiner, W., Waber, J. T.: *Theoret. chim. Acta (Berl.)* **21**, 235 (1971)
2. Fricke, B., Waber, J. T.: *Actinides Rev.* **1**, 433 (1971)
3. Malý, J., Hussonnois, M.: Report IPN-RC 72.02 Orsay, July 1972, *Theoret. chim. Acta (Berl.)* **28**, 363 (1973)

4. Rosen, A., Lindgren, I.: Phys. Rev. **176**, 114 (1968)
5. Swirles, B.: Proc. Roy. Soc. (London) A **152**, 625 (1935)
6. Grant, I. P.: Proc. Roy. Soc. (London) A **262**, 555 (1961)
7. Koopmans, T.: Physica **1**, 104 (1933)
8. Mann, J. B., Johnson, W. R.: Phys. Rev. A **4**, 41 (1971)
9. Slater, J. C.: Phys. Rev. **81**, 385 (1951)
10. Gaspar, R.: Acta Physiol. Acad. Sci. Hung. **3**, 263 (1954)
11. Slater, J. C.: Quantum theory of atomic structure, Vol. 1. New York-Toronto-London: Mc Graw Hill 1960
12. Bearden, J. A., Burr, A. F.: Rev. Mod. Phys. **39**, 125 (1967)
13. Grant, I. P.: Advan. Phys. **19**, 747 (1970)
14. Moore, Ch. E.: Atomic energy levels, Nat. Bur. of Standards Circular N° 467 (U. S. Government Printing Office, Washington, P.C., 1949)
15. Moore, Ch. E.: NSRDS-NBS 34 Report, Nat. Stand. Ref. Data Ser., National Bureau of Standards, Washington 1970
16. Moore, Ch. E.: Appl. Opt. **2**, 665 (1963)
17. Smith, D. H., Hertel, G. R.: J. Chem. Phys. **51**, 3105 (1969)
18. Siegbahn, W., and others: ESCA, Atomic molecular and solid state structure studies by means of electron spectroscopy. Upsalla: Almquist and Wiksells, 1967
19. Cunningham, B. B.: Ref. 140 in Seaborg, G. T.: Ann. Rev. Nucl. Sci. **18**, 53 (1968)
20. David, F.: Report IPN RC- 71.06, Orsay, October 1971
21. Keller, O. L., Burnett, J. L., Carlson, T. A., Nestor, C. W.: J. Phys. Chem. **74**, 1127 (1970)
22. Desclaux, J. P., Moser, C. M., Verhagen, G.: J. Phys. B. Atom. Molec. Phys. **4**, 296 (1971)

Dr. J. Malý  
Institut de Physique Nucléaire  
Division de Radiochimie  
Université de Paris XI  
Centre d'Orsay  
18 F-91406 Orsay, France



# The Influence of Solvation on $\pi^* \leftarrow n$ and $\pi^* \leftarrow \pi$ Transition Energies in Molecules Containing Aza or Carbonyl Groups

P. Cremaschi, A. Gamba and M. Simonetta

C.N.R. Centre and Physical Chemistry Institute of the University of Milano, I-20133 Milano, Italy

Received April 16, 1973

The influence of solvation on  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies of formaldehyde, acetaldehyde, acetone, pyridine and 1,2-, 1,3-, 1,4-diazabenzene has been investigated through CNDO calculations.

A static solvation model which distinguishes a) molecules directly involved in hydrogen bonding with solute, b) the layer of molecules in contact with the solute molecule and c) the main of molecules farther from solute, is presented.

Blue and red shifts due to solvent effects are correctly predicted by calculations for each model.

**Key words:** Solvent Shifts of Electronic Transitions.

## 1. Introduction

At present it seems that CNDO—MO theory [1] offers the most promising tool for the study of the electronic properties of large molecular systems; the same method has also been extensively and successfully used to obtain a "semi-quantitative" description of hydrogen bonded structures [2].

This method has been recently employed to study solvent effects on the activation energy of  $\text{CH}_3\text{F} + \text{F}^-$  reaction [3].

In the present paper we intend to study the influence of the solvent on electronic spectra by using CNDO/2 method in its original parametrization [4].

An analogous problem has been previously investigated by Hoffman *and coworkers* [5] who focused their attention on the influence of hydrogen bond formation on different observables, including  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies, by using a more empirical approach namely the Extended Huckel Method.

The most convenient choice of solutes should meet the following two requirements: a simple geometry to favour the formulation of a model of solvation, and a well established location of the  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transitions in gas phase and solution spectra. With this in view formaldehyde, acetaldehyde, acetone, pyridine and diazabenzene have been considered.

The solvent should be made of small strongly polar molecules in order to save computation time and to produce remarkable shifts of the bands with respect

to those obtained in gas phase or in apolar solvents. Water meets these requirements. Moreover by using the same solvent as in Ref. [5] an interesting comparison between results obtained by CNDO and EH methods is possible.

## 2. Calculations

The geometries and the ground state properties of the solute molecules and of water have been optimized by the energy minimizing procedure due to Powell [6], where the energy was calculated by CNDO/2 method [4].

Excited-state properties have been obtained, using the same geometry as in the ground state, by CNDO/CI method [7], where the electronic repulsion integrals were evaluated by the Pariser and Parr technique [8]. The configuration interaction included the 30 mono-excited configurations of lower energy.

Calculated spectra of isolated molecules were compared with spectra measured in vacuo. When spectra obtained in solution were considered, calculations were performed on a model for the solvated molecule. The size of calculation and the necessity of preserving the point symmetry of the solute molecule do not allow us to consider an extended solvation process. We have tried to describe solvent effects by adding to the solute molecule a few solvent molecules, namely by hydrating to different extents the proper positions of the solute. No more than five molecules of water have been located around each position. The packing with the solvent was studied for different values of the intermolecular distance  $q$  (values ranged from 1.4–2.4 Å). This procedure was preferred to a minimizing process with respect to the intersystem distance, as it is well known that the CNDO method systematically overestimates bonding properties. In the description of the geometries of the solute-solvent systems most of the geometrical parameters found for the isolated molecules have been used. However for each intersystem distance some geometrical parameters, involving the site of interaction, which will be specified in each case, have been optimized. As an example in the case of the molecule of water, the bond angle and the O—H bond distance not involved in the intermolecular bond have been kept equal to the values optimized for the isolated molecule, while the O—H distance of the bond involved in the solvation procedure has been let vary.

Once the geometries of minimum energy of the ground state were obtained for each  $q$  value, the spectrum was calculated by the CI procedure [7].

## 3. Results

The topology of the isolated molecules of solutes, that is formaldehyde, acetaldehyde, acetone, pyridine and 1,2-, 1,3- and 1,4-diazabenzenes are given in Fig. 1, where the optimized geometrical parameters are also reported. In the corresponding Table 1 a comparison between experimental and calculated geometries is shown. In the same table the optimized geometry of the isolated molecule of water is reported. In the case of azines all C—H bonds were assumed in the ring-plane bisecting ring-angles, with equal bond-lengths.

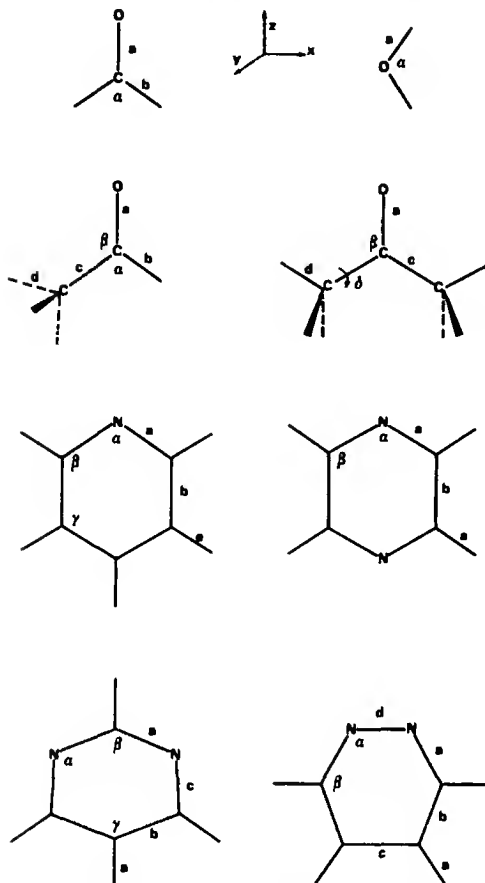


Fig. 1. Geometries and optimized parameters for isolated systems

## a) Formaldehyde

The solvent effect on formaldehyde has been studied by considering a progressive solvation, that is several packings in which the formaldehyde molecule is bound up to six units of solvent. The more realistic situations are illustrated in Fig. 2, where the geometrical parameters which have been optimized are shown. The minimizing process has been performed for six different values of the  $O \cdots H$  ( $\rho$ ) distance, which was varied from 1.4–2.4 Å in steps of 0.2 Å<sup>1</sup>. In the cases (c) and (d), where four molecules of solvent are considered, a higher stabilization is observed for (d) structure; this difference of stabilization between the two structures reaches the largest value of 24 kcal/mole for  $\rho = 2.0$  Å. If the methylene group is also solvated (on the whole six molecules of hydration), the difference in energy stabilization decreases to about 1–2 kcal/mole for all  $\rho$  values.

<sup>1</sup>Optimized geometries are not reported, but are available on request.

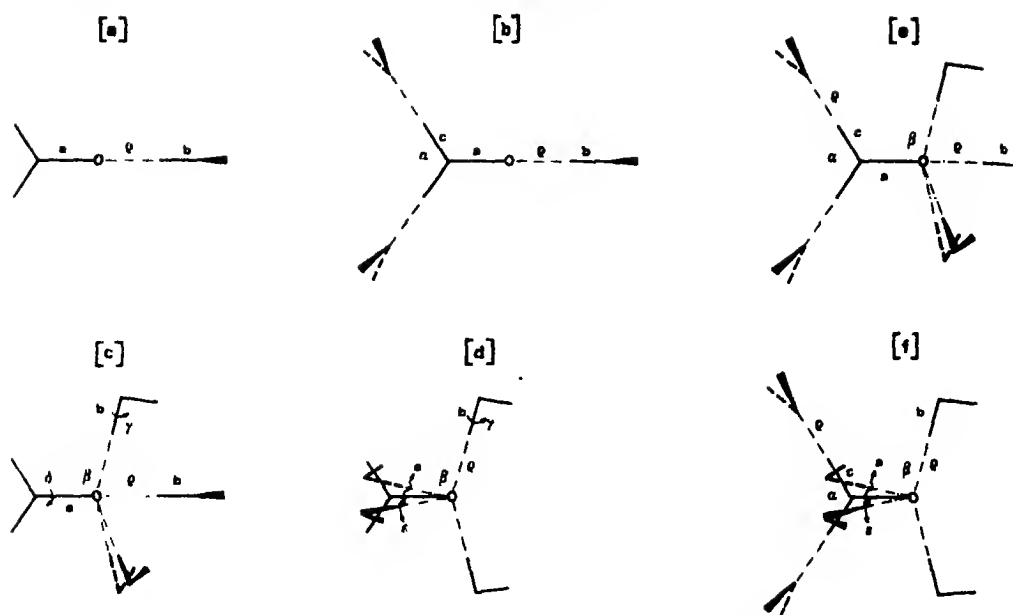


Fig. 2. Solvation model of formaldehyde with 1, 3, 4 and 6 water molecules

Table 1. Experimental and optimized geometries of the isolated molecules\*

Molecules	Geometrical parameters							
	a	b	c	d	e	$\alpha$	$\beta$	$\gamma$
HCHO	1.247 <i>1.2078<sup>b</sup></i>	1.114 <i>1.1161</i>				115.7 <i>116.5</i>		
CH <sub>3</sub> CHO	1.261 <i>1.2155<sup>c</sup></i>	1.123 <i>1.114</i>	1.443 <i>1.5005</i>	1.119 <i>1.0861</i>		111.2 <i>117.5</i>	126.5 <i>123.9</i>	
CH <sub>3</sub> COCH <sub>3</sub>	1.274 <i>1.240<sup>c</sup></i>		1.455 <i>1.52</i>	1.119 <i>1.09</i>			120.2 <i>118.5</i>	111.6 <i>109.28</i>
C <sub>6</sub> H <sub>5</sub> N	1.342 <i>1.3402<sup>d</sup></i>	1.382 <i>1.3945</i>			1.116 <i>o-1.0843</i> <i>m-1.0805</i> <i>p-1.0773</i>	116.3 <i>116.83</i>	124.9 <i>123.88</i>	117.5 <i>118.53</i>
1,4-C <sub>6</sub> H <sub>4</sub> N <sub>2</sub>	1.344 <i>1.334<sup>e</sup></i>	1.381 <i>1.378</i>			1.121 <i>1.05</i>	112.0 <i>115.8</i>		
1,3-C <sub>6</sub> H <sub>4</sub> N <sub>2</sub>	1.342 <i>1.335<sup>f</sup></i>	1.380 <i>1.395</i>	1.343 <i>1.355</i>		1.119	118.5 <i>115.1</i>	130.2 <i>128.2</i>	115.9 <i>116.3</i>
1,2-C <sub>6</sub> H <sub>4</sub> N <sub>2</sub>	1.351	1.374	1.384	1.280	1.116	119.5 <i>119.0<sup>f</sup></i>	124.0 <i>123.68</i>	
H <sub>2</sub> O	1.029 <i>0.9571<sup>c</sup></i>					104.5 <i>104.5</i>		

\* All the distances are in Å and the angles in °. The upper number refers to the calculated value and the lower (in italics) refers to the experimental one.

<sup>b</sup> See Ref. [9]. <sup>c</sup> See Ref. [10]. <sup>d</sup> See Ref. [11]. <sup>e</sup> See Ref. [12]. <sup>f</sup> See Ref. [13].



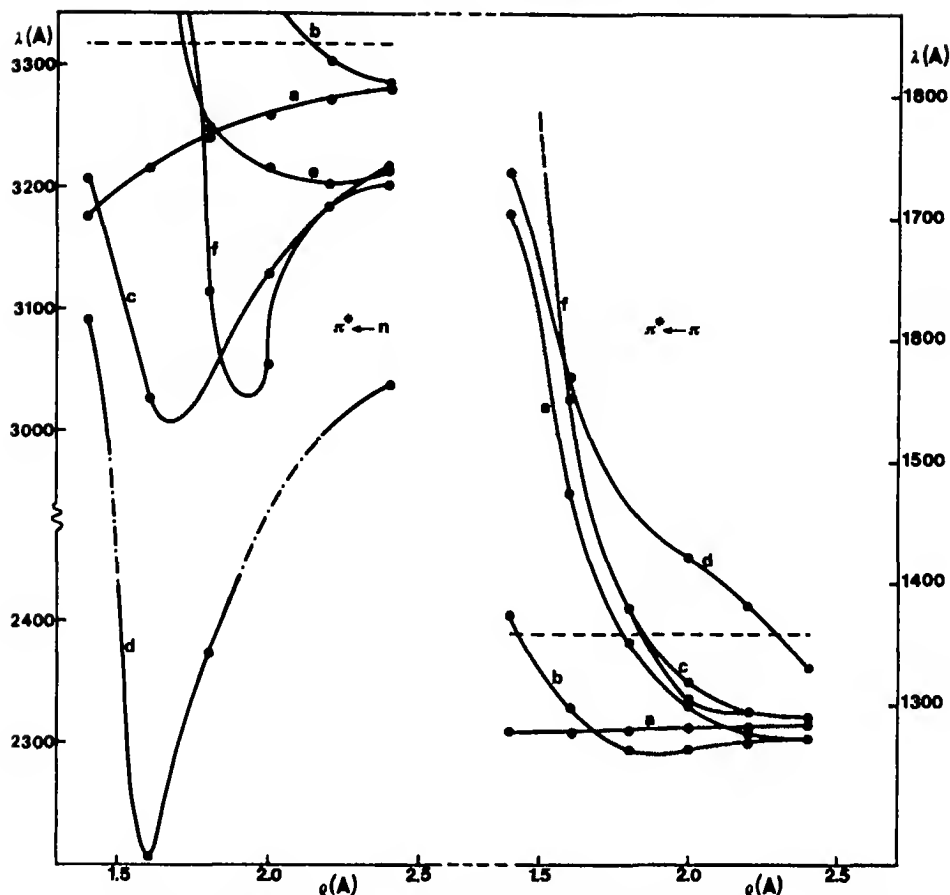


Fig. 3. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $q$  of the different solvation models for formaldehyde. Dot line represents transition energy for isolated molecule

The electronic spectrum of formaldehyde has been calculated by considering the molecule in the solvated situations illustrated in Fig. 2. The packings (e) and (f) correspond to (c) and (d) packings respectively, when the methylene group was hydrated. It is known from experiment that  $\pi^* \leftarrow n$  transition of formaldehyde falls at 3538 Å in vacuo [14], and shifts to 2900 Å [15] in water. The trend of calculated  $\pi^* \leftarrow n$  transition in different packings of solvent, for several  $q$  values, is shown in Fig. 3.

A correct blue shift is predicted in the cases (a), (c) and (d) for all  $q$  values. In the case (b) the blue shift is found only for the highest values of  $q$ . In general it seems that the correct trend can be found when the hydration of C—O group is stronger than that of CH<sub>2</sub> group.

The predicted red shift for  $\pi^* \leftarrow \pi$  transition [16, 17] is obtained by our calculations for the higher degrees of solvation and, in general, for shorter values of  $q$ . The case (d), where the red shift is found for all  $q$  values, is the exception.

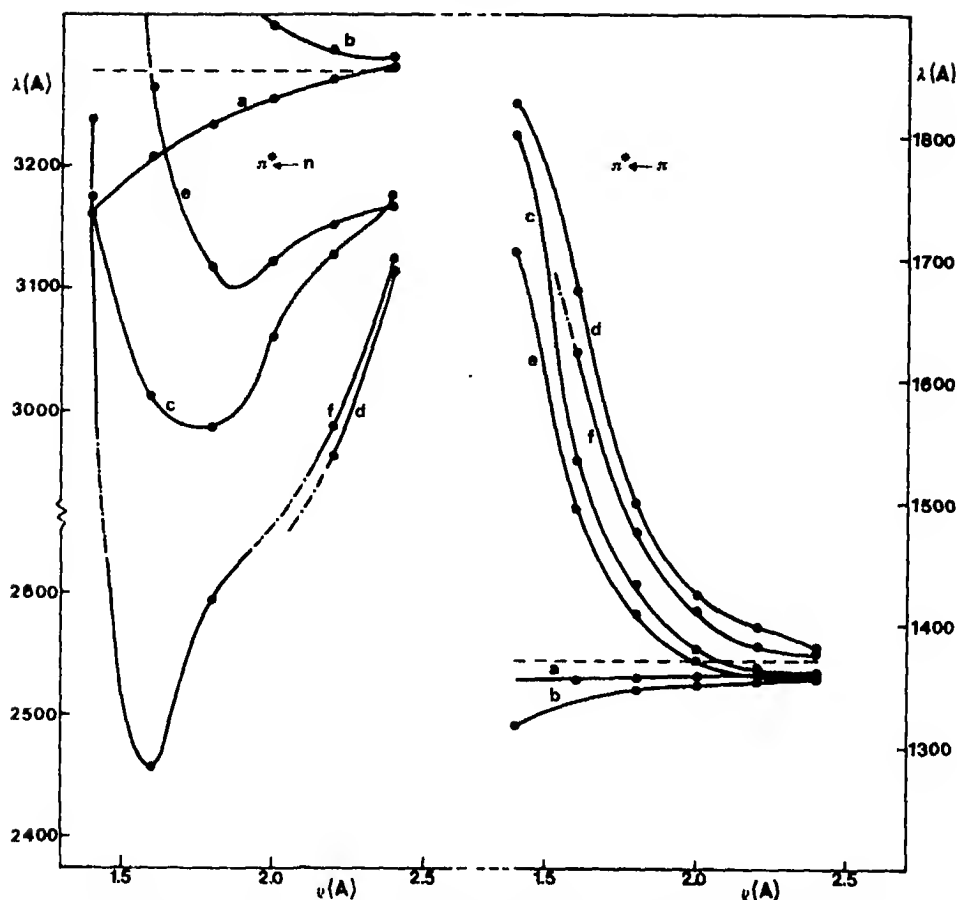


Fig. 4 Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $\rho$  of the different solvation models for acetaldehyde. Dot line represents transition energy for isolated molecule

#### b) Acetaldehyde and Acetone

For both molecules several solvation models were considered, using one, four or five molecules of solvent in the case of acetaldehyde and one or four molecules of solvent in the case of acetone. On the basis of the experience of previous calculations on  $\text{CH}_3\text{F}$  [3], the methyl group was never hydrated. In view of this the solvent models are the same as those considered for formaldehyde.

The solvents shifts of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  bands for the two molecules as a function of hydration and  $\rho$ -distance are shown in Figs. 4 and 5. The  $\pi^* \leftarrow n$  transition of acetaldehyde, which falls at 3390 Å in vacuo [18], shifts to 2740 Å in water [17]. For acetone the  $\pi^* \leftarrow n$  transition located at 2762 Å in vacuo [17], occurs at 2640 Å in water [19]. MO calculations in both cases give a correct interpretation of the phenomenon, showing a pronounced increase of the blue shift as more molecules of solvent are put around the carbonyl group. Moreover

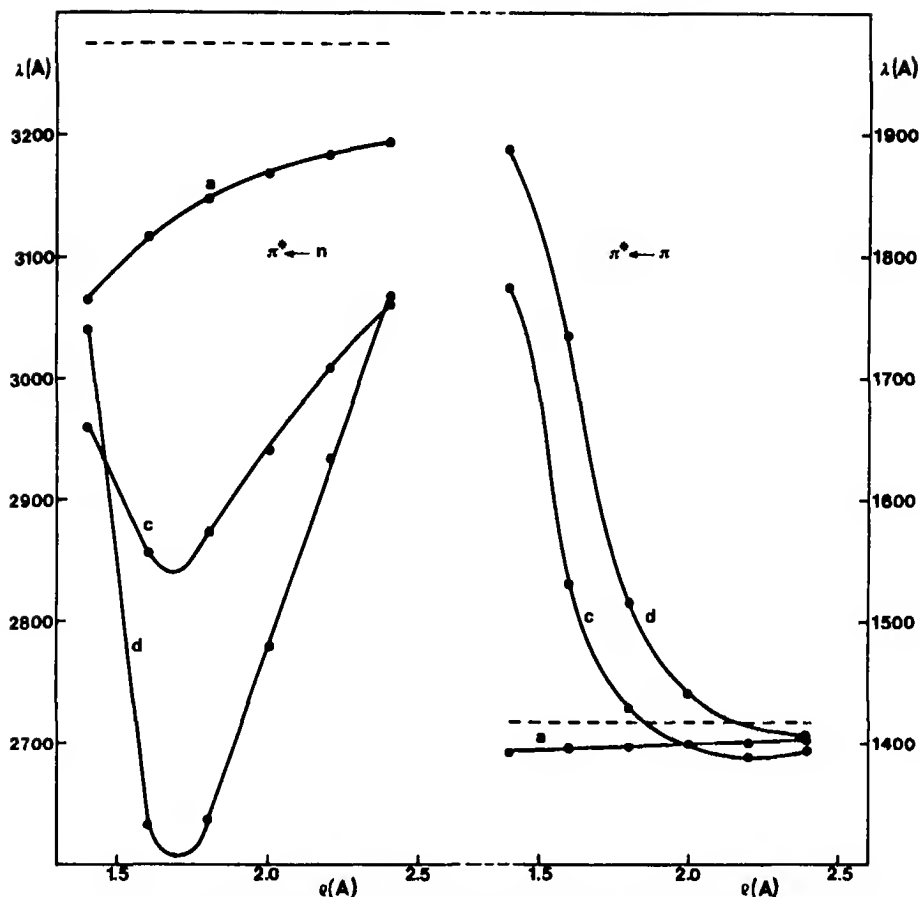


Fig. 5. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $q$  of the different solvation models for acetone. Dot line represents transition energy for isolated molecule

a red shift is predicted by our calculations in water for  $\pi^* \leftarrow \pi$  transitions of both molecules only when a sufficient number of molecules of solvent are considered. In conclusion it seems that for these two molecules the solvent effect on electronic transitions is well interpreted by CNDO calculations only when at least four molecules are around the solute.

### c) Pyridine

For water-pyridine solution several models were investigated by considering up to six units of solvents for each solute molecule. When one molecule of water has been used to represent the solvent effect, only the configuration in which one of the protons of the solvent approaches the nitrogen lone pair has been considered, having in mind the results obtained by Hoffmann and coworkers (see Fig. 3 in Ref. [5]). The models assumed when more molecules of water were considered, are shown in Fig. 6, where the optimized geometrical parameters are also reported.

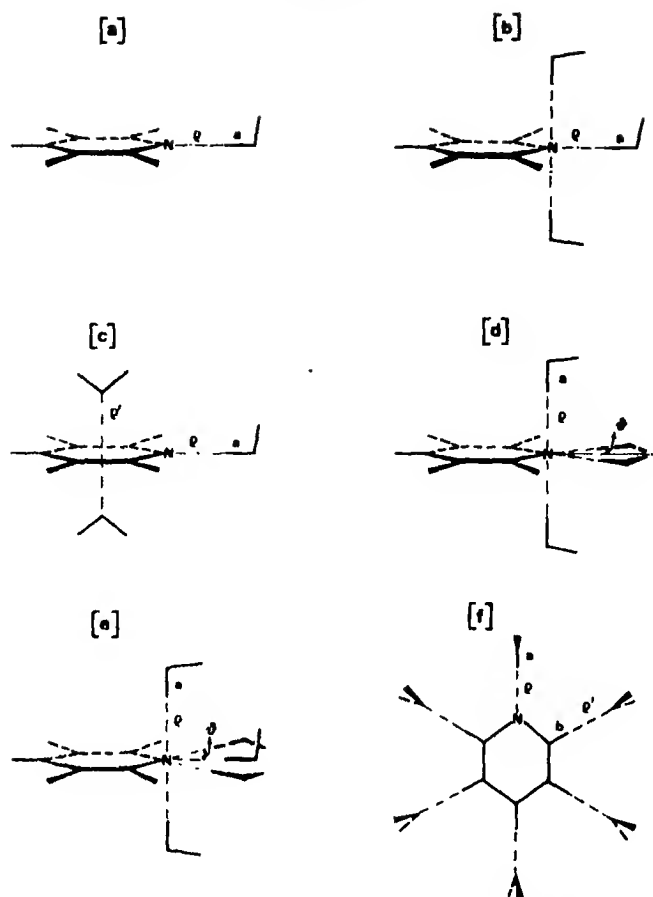


Fig. 6. Solvation model for pyridine with 1-6 water molecules

The calculated solvent shifts of electronic bands of pyridine in different solvent situations are shown in Fig. 7 as a function of the  $q$  distance. A blue shift of about 300–400 Å is expected for  $\pi^* \leftarrow n$  transition in water [20, 21]. Our results give the correct answer for all the considered models and the quantitative agreement with experiment is the better the greater the number of molecules of water around the nitrogen. Likewise the observed red shift of the  $p$  band of pyridine, which was found at 2500 Å in vacuo [22] and at 2530 Å [21] (or 2570 Å [20]) in water, is correctly predicted by calculations for all  $q$  values and in this case the agreement is excellent also on a quantitative basis, with the exception of (c) model, where the red shift is overestimated.

#### d) Diazabenzene

Solvation models with two, six, eight and ten molecules of water were studied for 1,4-, 1,3- and 1,2-diazabenzene. Owing to its topology in the case of 1,2-isomer only the models with two and six molecules of water have been

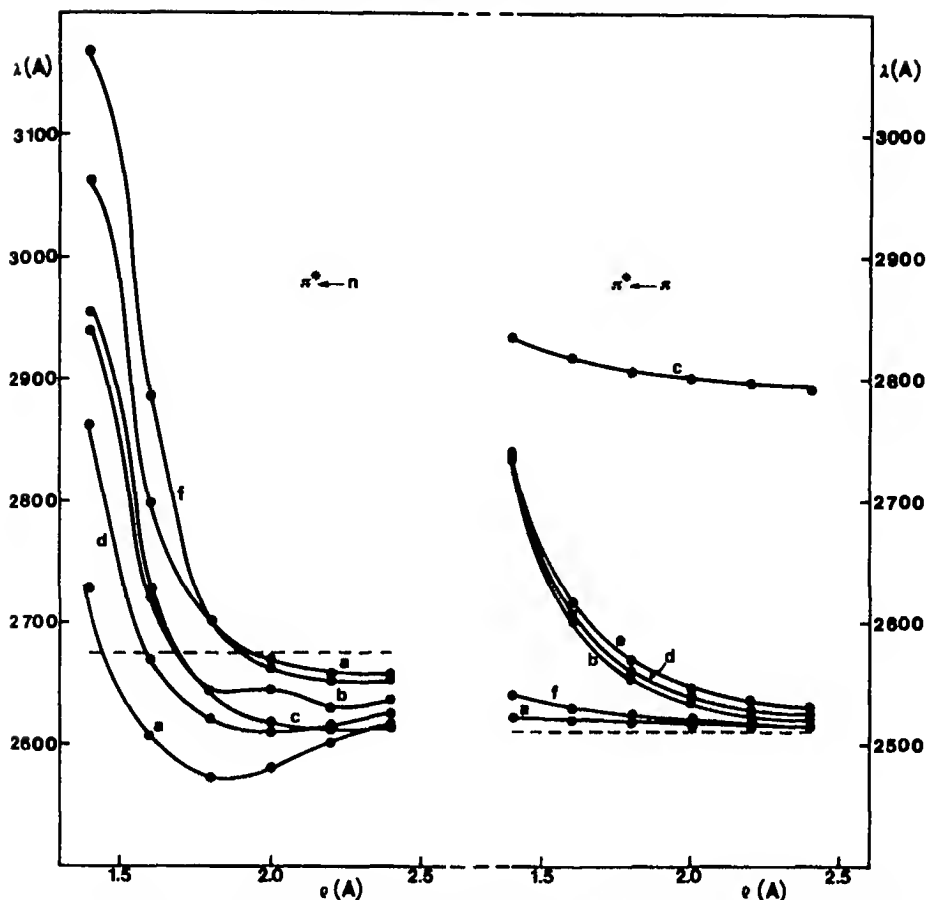


Fig. 7. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $\rho$  of the different solvation models for pyridine. Dot line represents transition energy for isolated molecule

considered. The structures of hydrated pyridine were used as reference models, keeping in mind that in diazabenzenes two nitrogens are present.

As found for the former molecules the solvent shifts for  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  are well predicted by calculations. Now the agreement with experiment for  $\pi^* \leftarrow n$  is even better, while in the case of  $\pi^* \leftarrow \pi$  transitions the red shift, which is experimentally found to be very low in all cases and practically zero for 1,2-diazabenzene, is slightly overestimated. The trend of transition energies for the three molecules in different solvent structures, for different values of  $\rho$ , is shown in Figs. 8, 9 and 10.

#### 4. Discussion

In a static model of solvation three types of solvent molecules can be distinguished: A) molecules directly involved in hydrogen bonding to the solute; B) the layer of molecules in contact with the solute molecule, that is

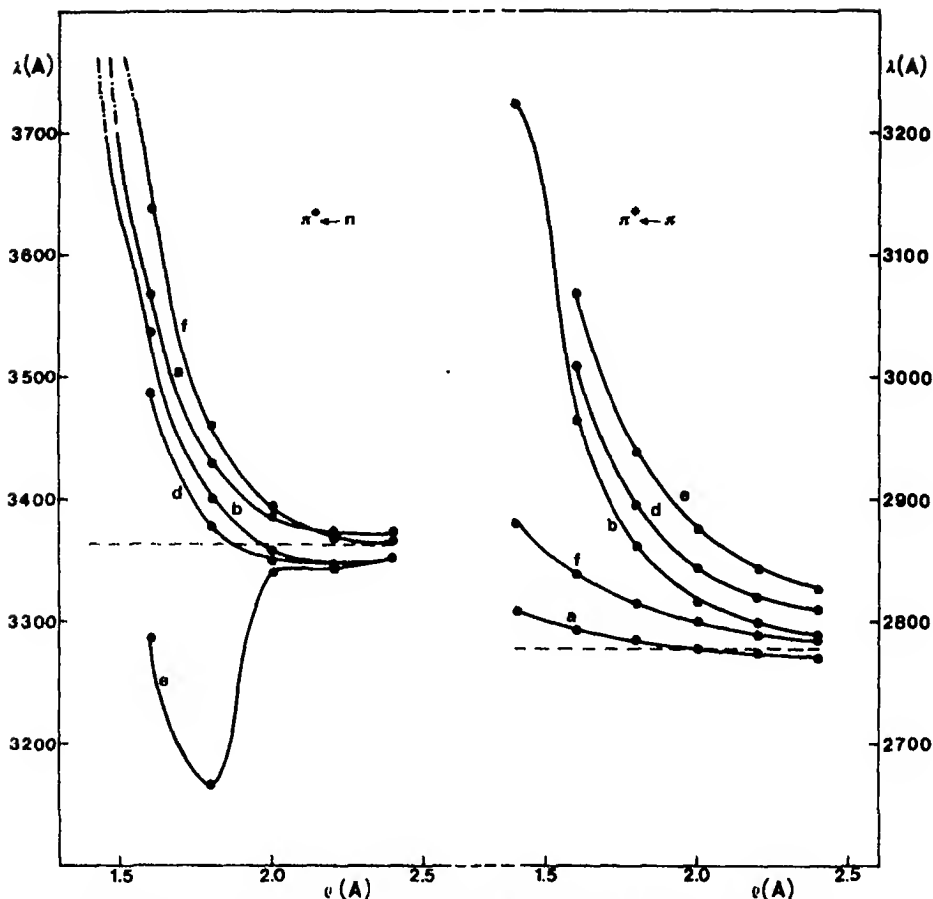


Fig. 8. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $q$  of the different solvation models for 1, 4-diazabenzene. Dot line represents transition energy for isolated molecule.

directly interacting with it, but without forming hydrogen bonds, and C) the surrounding molecules farther from the solute.

In our theoretical model as few as possible water molecules of type B have been considered, and the effect of the molecules of type C has been considered negligible. The geometries of the models for the different solutes are shown in Fig. 1 and have been previously discussed. By inspection of Table 1 a very satisfactory agreement between geometries calculated by CNDO method and those obtained by experiment for all isolated molecules can be verified.

When only one molecule of water is considered, as is usually done when studying hydrogen bonding [5], a stabilization energy ranging between 7 and 10 kcal/mole, is found. This order of magnitude reproduces well the estimated value for a hydrogen bond of such type [17]. The maximum stabilization energy for the different molecules, interacting with one unit of water (*a* case) are collected in Table 2, where the corresponding  $q$  distances are also shown.

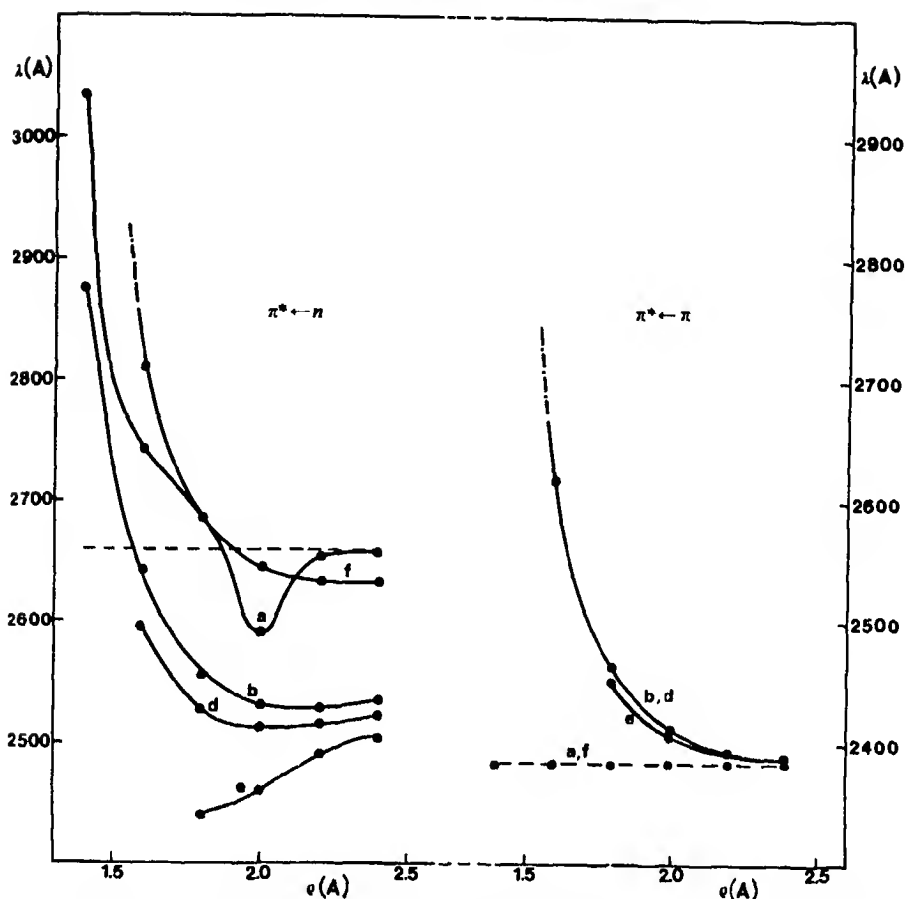


Fig. 9. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $\rho$  of the different solvation models for 1, 3-diazabenzene. Dot line represents transition energy for isolated molecule

For the considered molecules the distance that minimizes the energy is 1.6 Å. The exception is acetone – water system for which a  $\rho$  value of 1.4 Å is obtained. CNDO results for pyridine, that is a stabilization energy of  $-10.7$  kcal/mole for an equilibrium distance between oxygen and nitrogen atoms of 2.65 Å, can be compared with  $-2.3$  kcal/mole obtained by EH method with an equilibrium distance of 2.76 Å [5].

When more units of solvent are packed around the solute, the solute – solvent mean distance rises to values between 1.8 and 2.0 Å and the corresponding stabilization energy of solvation model is 30–50 kcal/mole. The relevant numerical data are reported in Table 2.

The solvent effect on transition energies has been studied considering different packings at several solvent – solute distances. The relevant data are shown in Figs. 3–5, 7–10. When a sufficient number of water units are packed around the solute a satisfactory interpretation of the solvation effect can be obtained by CNDO calculations if a favourable  $\rho$  value is chosen. The

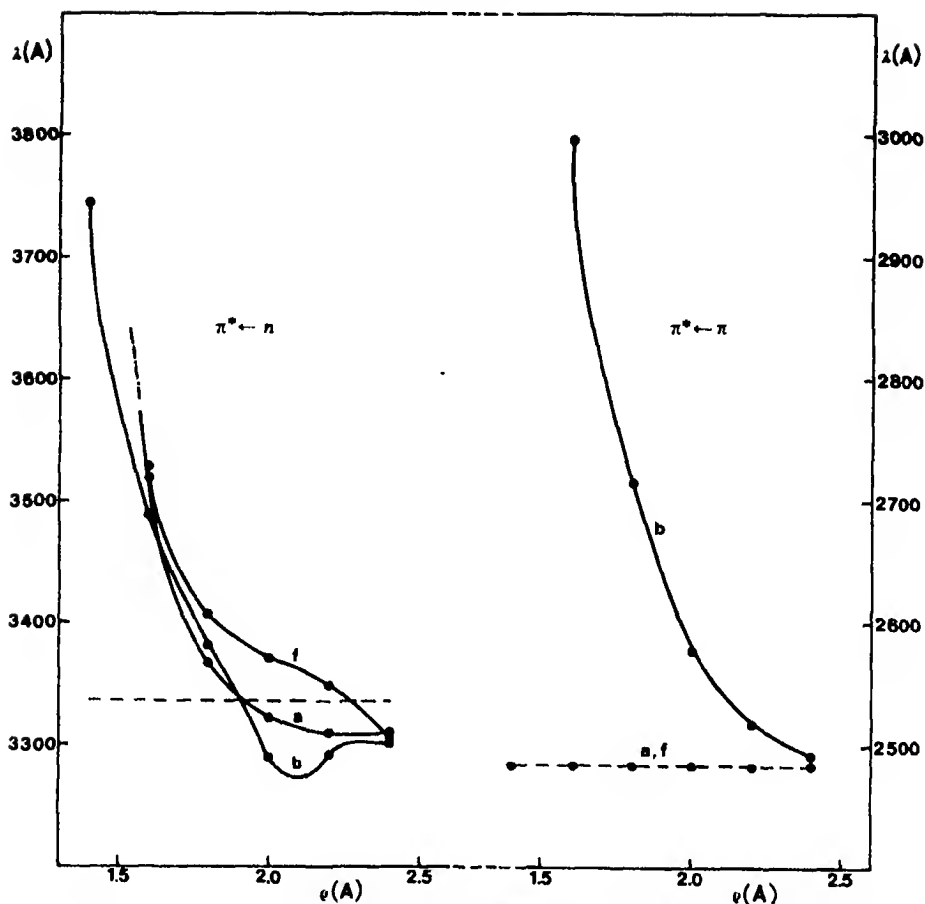


Fig. 10. Trends of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transition energies vs  $\rho$  of the different solvation models for 1, 2-diazabenzene. Dot line represents transition energy for isolated molecule

situation is illustrated in Table 3 where a qualitative agreement between measured and calculated blue and red shifts for  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transitions respectively, can be observed for all molecules when a proper amount of solvation is included; the  $\rho$  values at which the best agreement is obtained are also shown.

In Table 4 oscillator strengths ( $f$ ) for  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transitions, calculated for isolated molecule and for the same hydrated models as presented in Table 3, are shown;  $f$  values for  $\pi^* \leftarrow n$  transitions of isolated molecules are always higher than the values obtained for solvated molecules, however very small  $f$  values are always obtained. On the contrary an increasing of  $f$  values for  $\pi^* \leftarrow \pi$  transitions is found for hydrated molecules compared to isolated ones. The few available experimental data confirm this finding, apart from the exception of the  $\pi^* \leftarrow \pi$  band of acetone, where a lowering of oscillator strength of  $\pi^* \leftarrow \pi$  band for isolated molecule respect to hydrated one is theoretically and experimentally found.



Table 2. Stabilization energies for several solvation models

Molecule	Model <sup>a</sup>	$q$ (Å)	$\Delta E$ (kcal/mole)
HCHO	a	1.6	7.09
	d	2.0	39.09
CH <sub>3</sub> CHO	a	1.6	7.59
	d	1.8	38.72
CH <sub>3</sub> COCH <sub>3</sub>	a	1.4	9.29
	d	2.0	31.12
C <sub>5</sub> H <sub>5</sub> N	a	1.6	10.67
	c	1.8	21.15
1,4-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	a	1.6	19.14
	c	1.8	41.73
1,3-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	a	1.6	20.46
	c	1.8	38.28
1,2-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	a	1.6	19.20
	b	2.0	37.84

<sup>a</sup> The models (a), (b), (d) and (c) refer to the corresponding structures in Figs. 2 and 6.

Table 3. Experimental and calculated blue and red shifts for  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transitions

Molecule and model	$q$ (Å)	$\pi^* \leftarrow n$ transition blue shift (nm)		$\pi^* \leftarrow \pi$ transition red shift (nm)	
		calc.	exp.	calc.	exp.
HCHO (d)	2.0	70.4	63.8	6.3	
CH <sub>3</sub> CHO (d)	1.8	68.1	65.0	5.4	
CH <sub>3</sub> COCH <sub>3</sub> (d)	2.0	49.3	12.2	2.6	
C <sub>5</sub> H <sub>5</sub> N (c)	1.8	9.8	30.0–34.0	5.9	3.0–7.0
1,4-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> (c)	1.8	19.7	17.5–23.9	16.3	3.9
1,3-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> (c)	1.8	21.9	18.8–24.9	6.9	1.5
1,2-C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> (b)	2.0	4.5	25.0–35.6	23.3	~0

In the case of pyridine, for which the change in the electric dipole moment upon excitation to the lowest energy singlet  $\pi^* \leftarrow n$  excited state has been measured [26] to be  $\Delta\mu = 3.2$  D, a quantitative agreement with experiment has been found. The measured ground state dipole moment of pyridine is 2.19 D [27], consequently the excited state dipole moment is  $-1.0$  D. Our CNDO calculation predicts 2.94 D for the ground state dipole moment and  $-1.44$  D for excited state dipole, with a substantial decreasing of the moment upon excitation, as predicted by solvent effect theory when a blue shift is observed [28]. The calculated dipole moment for the excited state, associated with the first  $\pi^* \leftarrow \pi$  transition is 3.06 D, according to observed red shift. Moreover for the other diazabenzenes CNDO calculations predict for an observed blue or red shift, a corresponding lower or higher value of the dipole of the involved excited state, with respect to the ground state dipole moment. In the case of formaldehyde, acetaldehyde and acetone for a calculated blue shift of  $\pi^* \leftarrow n$  transition a decrease of the value of dipole moment of excited  $\pi^*$  state (with respect to the  $n$

Table 4. Calculated and observed oscillator strengths ( $f$ ) of  $\pi^* \leftarrow n$  and  $\pi^* \leftarrow \pi$  transitions

Molecule	$\pi^* \leftarrow n$		$\pi^* \leftarrow \pi$	
	in vacuo	in water	in vacuo	in water
HCHO		-	0.228	0.354
		-	0.57 <sup>a,h</sup>	-
CH <sub>3</sub> CHO			0.383	0.278
			0.20 <sup>b,h</sup>	
CH <sub>3</sub> COCH <sub>3</sub>		-	0.360	0.352
			0.16 <sup>c,h</sup>	0.00 <sup>d,h</sup>
C <sub>4</sub> H <sub>4</sub> N	0.005	0.003	0.083	0.100
				0.07 <sup>e,h</sup> - 0.06 <sup>f,h</sup>
1,4 -C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	0.009	0.003	0.206	0.249
	0.012 <sup>a</sup>	0.017 <sup>f,h</sup>	0.10 <sup>a</sup>	0.12 <sup>f,h</sup>
1,3 -C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	0.000	0.003	0.080	0.102
	0.0076 <sup>a</sup>	0.01 <sup>f,h</sup>	0.052 <sup>a</sup>	0.06 <sup>f,h</sup>
1,2 -C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>	0.012	0.011	0.071	0.080
	0.008 <sup>a</sup>	0.006 <sup>f,h</sup>	0.020 <sup>a</sup>	

<sup>a</sup> See Ref. [23]   <sup>b</sup> See Ref. [24]   <sup>c</sup> See Ref. [25]   <sup>d</sup> See Ref. [19]   <sup>e</sup> See Ref. [21]   <sup>f</sup> See Ref. [20]   <sup>g</sup> See Ref. [22]   <sup>h</sup> Calculated from  $\epsilon_{\max}$  through the relationship  $f = 2 \cdot 10^{-3} \cdot \epsilon_{\max}$ .

state dipole moment) is found. On the other hand for  $\pi^* \leftarrow \pi$  transitions, for which a correct red shift is obtained, the calculated dipole moments for excited  $\pi^*$  states associated with transitions are again lower than ground state dipole moments. However it is well known that not all frequency shifts are caused entirely by dipole interactions, but other types of interaction may give important or even dominant contributions in certain cases [28].

In conclusion we feel that the present hydration model gives an adequate description of solvent effect, keeping in mind the limits imposed by the complexity of the problem. It is noteworthy that the observed shift of transition energy is correctly predicted only when the  $\rho$  values are in the range of the experimental finding and a sufficient number of water molecules is put around the solute. This fact suggests that hydrogen bond model alone cannot explain observed trends of the UV spectra due to solvent effect. It is the first time to our knowledge that the blue and red shifts due to solvent effects can be correctly predicted by one calculation.

*Acknowledgements.* One of us (P.C.) thanks Professor C. Moser for hospitality at C.E.C.A.M. laboratories in Orsay.

## References

1. Pople, J.A., Beveridge, D.L.: Approximate molecular orbital theory. McGraw Hill 1970
2. Schuster, P.: Theoret. chim. Acta (Berl.) **19**, 212 (1970)  
Kollman, P.A., Allen, L.C.: Chem. Rev. **72**, 283 (1972)
3. Cremaschi, P., Gamba, A., Simonetta, M.: Theoret. chim. Acta (Berl.) **25**, 237 (1972)
4. Pople, J.A., Segal, G.A.: J. Chem. Phys. **44**, 3289 (1966)
5. Adam, W., Grimison, A., Hoffmann, R., de Hortiz, C.Z.: J. Am. Chem. Soc. **90**, 1509 (1968)

6. Powell, M. J. D.: *Computer J.* **7**, 303 (1965)  
Powell, M. J. D., VA04A: Minimum of function of several variables, Quantum Chemistry Exchange Program. Indiana University 1965
7. Del Bene, J. E., Jaffè, H. H.: *J. Chem. Phys.* **48**, 1807 (1968)  
Ellis, R. L., Kuehlenz, G., Jaffè, H. H.: *Theoret. chim. Acta (Berl.)* **26**, 131 (1972)
8. Pariser, R., Parr, R. G.: *J. Chem. Phys.* **21**, 767 (1953)
9. Tagaci, K., Hoka, T.: *J. Phys. Soc. Japan* **18**, 1174 (1963)
10. Tables of interatomic distances and configurations in molecules and ions. London: Spec Publ. n° 11, Scient. Ed. Sutton. The Chemical Society 1958
11. Hansen-Nygaard, B. L., Rastrup-Andersen, J.: *J. Mol. Spectr.* **2**, 54 (1958)
12. Wheatley, P. J.: *Acta Cryst.* **10**, 182 (1957)
13. Innes, K. K., Byrne, J. P., Ross, I. G.: *J. Mol. Spectr.* **22**, 125 (1967)
14. Brand, J. C. D.: *J. Chem. Soc.* 858 (1956)
15. Bercovici, T., King, J., Becker, R. S.: *J. Chem. Phys.* **56**, 3956 (1972)
16. Pimentel, G. C.: *J. Am. Chem. Soc.* **79**, 3323 (1957)
17. Pimentel, G. C., McClellan, A. L.: *The hydrogen bond*. W. H. Freeman 1960
18. Walsh, A. D.: *J. Chem. Soc. (London)* **1953**, 2318
19. Hayes, W. P., Timmons, C. J.: *Spectr. Acta* **21**, 529 (1965)
20. Mason, S. F.: *J. Chem. Soc. (London)* **1959**, 1240, 1247
21. Jaffè, H. H., Orchim, M.: *Theory and application of ultraviolet spectroscopy*. New York: John Wiley, 1962
22. Clevers, H. B., Platt, J. R.: Survey of vacuum U.V. spectra, Technical Report, Part One, 1953/54
23. Fleming, G., Anderson, M. M., Arrison, A. J., Pickett, L. W.: *J. Chem. Phys.* **30**, 351 (1959)
24. Lake, J. S., Arrison, A. J.: *J. Chem. Phys.* **30**, 361 (1959)
25. Barnes, E. F., Simpson, W. T.: *J. Chem. Phys.* **39**, 670 (1963)
26. Hochstrasser, R. M., Michaluk, J. W.: *J. Chem. Phys.* **55**, 4668 (1971)
27. Nelson Jr., R. D., Lide Jr., D. R., Maryott, A. A.: *Natl. Std. Ref. Data Ser., Natl. Bur. Std. (U.S.)* **10**, 29 (1967)
28. McRae, E. G.: *J. Phys. Chem.* **61**, 562 (1957)

Prof. Dr. M. Simonetta  
Università di Milano  
Istituto di Chimica Fisica  
Via Saldini, 50  
I-20133 Milano, Italy



# *Ab initio* Calculations of Core Electron Binding Energies and Shifts in Halomethanes

David B. Adams and David T. Clark

Department of Chemistry, University of Durham, South Road, Durham

Received March 27, 1973

*Ab initio* LCAO MO SCF calculations have been carried out to predict core electron binding energies and shifts in fluoro- and chloro-methanes. The quality of the calculations ranges from a better than double zeta basis set to minimal STO (3 G) basis set. Predictions of binding energies and shifts are made using Koopmans' theorem, hole state calculations and equivalent cores calculations. Using a flexible basis set there is very little difference in the prediction of shifts by these three methods but for minimal basis set calculations the equivalent cores calculations give the best results.

**Key words:** Core holes – Equivalent cores – Koopmans' theorem – ESCA shifts.

## 1. Introduction

The accurate measurement of molecular core binding energies using ESCA has stimulated an interest in non-empirical calculations of shifts in core electron binding energies [1, 2]. Three approaches have commonly been used within the Hartree Fock formalism.

1. *Koopmans' Theorem* [3] which equates binding energies to the negative of computed orbital energies.

2. *Hole States*-binding energies are computed as energy differences between the neutral molecule and hole states formed by the removal of core electrons [4].

3. *Equivalent Cores Method*-Shifts in core binding energies are computed from heats of reaction for the isodesmic processes involved in the thermodynamic equivalent cores model [5, 6].

Implicit in all of these approaches is the neglect (or assumed self cancellation) of correlation energy changes. For core levels, however, correlation energy corrections to core electron binding energy shifts are essentially atomic in nature (since core orbitals are so localized) and it seems clear both from the success of these three models and from direct calculations [7] that correlation energy corrections are essentially constant for a given core level.

Koopmans' theorem predictions of shifts are expected to be basis set dependent and even for a large basis set at the Hartree Fock limit electronic relaxation is neglected. Thus, unless the electronic relaxation energy is constant or varies in a regular manner for a particular series of molecules then Koopmans' theorem is not expected to yield a quantitative description of shifts in core binding energies. The core hole state calculations take into account electronic

relaxation but may give convergence difficulties in the SCF procedure. When there is more than one equivalent centre in a molecule the question of localized versus non-localized hole states presents computational problems. The available evidence from direct calculations [7], from the success of the equivalent cores model (implicit in which is the concept of localized core holes) and the observation of certain shake-up transitions (formally forbidden if the hole state is delocalized) [8] is compelling in favour of the description of core hole states in such systems as being localized on the time scale of the ESCA experiment. However the theoretical treatment of such states is much more difficult than for delocalized hole states. In investigating the effect of electronic relaxation as a function of electronic environment and basis set therefore it is convenient to avoid such problems by considering a series of molecules where there is a unique hole state. The equivalent cores method depends on the energy of core exchange (in a series of isodesmic reactions) remaining constant and the calculated value for this energy may well depend on the basis set used.

While independent applications of these three approaches to the discussion of core binding energy shifts have been made (cf. Refs. [1, 2, 4, 6, 9]) no detailed comparisons of these methods and their dependence on basis sets have been carried out. It is the purpose of this paper to investigate in detail each of these methods to obtain information on their basis set dependencies, relaxation energies and the validity of the constancy of the energy of core exchange as a function of electronic environment. The systems studied are fluoro- and chloromethanes for which the experimental data are well documented [1, 10, 11].

## 2. Calculations

*Ab initio* LCAO MO SCF calculations on the molecules  $\text{CH}_{4-n}\text{F}_n$  ( $n = 0-4$ ),  $\text{CH}_3\text{Cl}$  and  $\text{CH}_2\text{Cl}_2$  together with the isoelectronic series  $\text{NH}_{4-n}\text{F}_n^+$ ,  $\text{NH}_3\text{Cl}^+$  and  $\text{NH}_2\text{Cl}_2^+$  were carried out using a better than double zeta basis set of optimised gaussian functions [12]. These consisted of 4s contracted to 3s for hydrogen (scale factor 1.2) and 9s, 5p contracted to 5s, 3p for carbon, nitrogen and fluorine. A 12s, 9p basis set was used for chlorine [13] and this was contracted to 7s, 5p according to the principles outlined by Dunning [12]. For ease of reference this basis set will be referred to later as "the large basis set". These calculations, except for  $\text{CH}_2\text{Cl}_2$  and  $\text{NH}_2\text{Cl}_2^+$  were performed using the IBMOL V LCAO MO SCF programme [14], all other calculations reported here were performed using the ATMOL group of programmes [15]. The programmes were implemented on an IBM 360/195.

Calculations on the series  $\text{CH}_{4-n}\text{F}_n$ ,  $\text{NH}_{4-n}\text{F}_n^+$  and  $^*\text{CH}_{4-n}\text{F}_n^+$  (where \* indicates a vacancy in the  $\text{C}_{1s}$  shell) were carried out using the following smaller basis sets.

1. The core orbitals were represented by four contracted gaussians and the valence orbitals, including  $\text{H}_{1s}$  (scale factor 1.2) were represented by four gaussian functions contracted to groups of 3 and 1 thus allowing a more flexible description of the valence orbitals. (STO 4.31 G basis set).

2. Each orbital was represented by three contracted gaussian functions with a 1.2 scale factor for the  $\text{H}_{1s}$  (STO 3 G basis set).

The exponents and coefficients used for these two basis sets were those obtained by Stewart [16], from a least squares fit of gaussian functions to Clementi's STO SCF atomic orbitals [17].

### 3. Results

#### *Koopmans' Theorem*

The Koopmans' theorem prediction of the binding energies and the shifts are shown in Table 1 together with the experimental values. The accuracy with which shifts in  $C_{1s}$  core binding energies are predicted, as expected, increases with increased flexibility of the basis set (Fig. 1), but even the large basis set overestimates the shift between  $CH_4$  and  $CF_4$  by  $\sim 22\%$ . Koopmans' theorem neglects electronic relaxation on ionization, however this is not a reasonable assumption. Gelius and Siegbahn [18] have divided the molecular electronic reorganization energy from atom A,  $E_A^{reorg}(\text{mol})$ , into two terms

$$E_A^{reorg}(\text{mol}) = E_A^{contr} + E_A^{flow}$$

where the first term is the reorganization energy gained by the contraction of the local charge distribution around nucleus A and is essentially atomic. The second term represents the redistribution of electron density in the remainder of the molecule. Using the differences between the calculated binding energies using the negative of the Hartree Fock orbital energies (Koopmans' theorem) and the differences in the total energies of the atom and ion Gelius and Siegbahn have estimated the atomic reorganization energy for the  $1s$  ionization of carbon to be 13.7 eV [18]. This value accounts for most of the difference between the experimental binding energies and the Koopmans' theorem values

Table 1. Koopmans' theorem predictions

Molecule	3 G BE	Shift	4.31 G BE	Shift	Large basis BE	Shift	Experimental <sup>a</sup> BE	Shift
$C_{1s}$ Shifts and binding energies (eV)								
$CH_4$	305.43	0.0	304.35	0.0	304.95	0.0	290.7	0.0
$CH_3F$	309.64	4.21	307.43	3.08	307.75	2.80	293.5	2.8
$CH_2F_2$	313.90	8.47	310.82	6.47	310.81	5.86	296.3	5.6
$CHF_3$	318.25	12.81	314.40	10.05	314.08	9.13	299.0	8.3
$CF_4$	322.69	17.26	317.96	13.61	317.38	12.43	301.7	11.0
$CH_3Cl$					307.49	2.54	292.3	1.6
$CH_2Cl_2$	—	—	—	—	309.77	4.82	293.9	3.1
$F_{1s}$ Shifts and binding energies (eV)								
$CH_3F$	704.50	0.0	713.24	0.0	714.90	0.0	692.4	0.0
$CH_2F_2$	705.70	1.20	714.41	1.17	716.13	1.23	693.1	0.7
$CHF_3$	706.91	2.41	715.60	2.36	717.31	2.41	694.1	1.7
$CF_4$	708.17	3.67	716.76	3.52	718.46	3.56	695.0	2.6

<sup>a</sup> See Refs. [2, 10, 11].

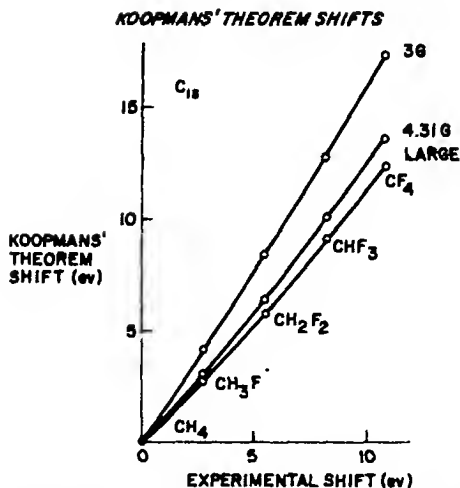
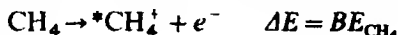


Fig. 1. Plot of Koopmans' theorem shifts (w.r.t. CH<sub>4</sub>) versus experimental shifts for the fluoromethanes as a function of basis set

in the cases of the 4.31 G and the large basis set calculations while the differences for the 3 G calculations are slightly larger. The estimate of a reorganization energy of 22.0 (or 22.1 employing a relativistic calculation) [18] for  $F_{1s}$  ionization accounts for most of the observed difference in the large basis set calculations but slightly overestimates the difference in the case of the 4.31 G calculations and grossly overestimates the difference in the case of the 3 G calculations. We attribute these differences to the poorer descriptions of the system given by the smaller basis sets. However the fact that with an improved basis set the shifts are well described by Koopmans' theorem suggests that reorganization energy differences contribute to only a minor extent for these closely related molecules. This will be discussed in more detail in a later section.

### Hole State Calculations<sup>1</sup>

The binding energy of a core electron e.g. in methane is the energy for the process



where \* indicates a vacancy in a core level, carbon 1s in this case. Since the photoionization process is rapid compared to nuclear motion the core hole states were taken to have geometries identical to the parent molecule and the total energies for 3 G and 4.31 G calculations on the series CH<sub>4-n</sub>F<sub>n</sub> and  ${}^*\text{CH}_{4-n}\text{F}_n^+$  are included in Table 2. The calculated binding energies and shifts

<sup>1</sup> For the hole states RHF calculations corresponding to the appropriate *locked* configurations have been carried out. There is no absolute guarantee that variational upper bounds to the true total energies for the ions are obtained since the computed hole states are not necessarily orthogonal to all lower energy states of the same symmetry. This could introduce errors of both a systematic and/or non-systematic nature. The results however would indicate that these considerations have not been encountered in this work.

Cf. Gianturco, F.A., Cuidotti, G., Chem. Phys. Letters 9, 539 (1971).



Table 2. Total energies (eV)

n	$\text{CH}_4-\text{F}_n$		$^*\text{CH}_4-\text{F}_n^+$		$\text{NH}_4-\text{F}_n^+$	
	3 G	4.31 G	Large	3 G	4.31 G	Large
0	-1078.1863	-	-1089.0380	-	-1514.4967	-1531.2662
1	-3720.4685	-	-3766.3545	-	-4153.8329	-4205.5299
2	-6363.0474	-	-6443.9813	-	-6793.0207	-6879.8908
3	-9005.7991	-	-9121.6151	-	-9431.9449	-9554.1786
4	-11648.5597	-	-11798.8994	-	-12070.5013	-12228.2038
			$\text{CH}_3\text{Cl}$	-11852.3984		$\text{NH}_3\text{Cl}^+$
			$\text{CH}_2\text{Cl}_2$	-13580.5603		$\text{NH}_2\text{Cl}_2^+$
				-26067.4141		-14023.0044
						-26508.2375

Table 3. Hole state calculations

	3 G		4.31 G		Experimental <sup>a</sup>	
	BE	Shift	BE	Shift	BE	Shift
$\text{C}_{1s}$ binding energies and shifts (eV)						
$\text{CH}_4$	297.51	0.0	292.93	0.0	290.7	0.0
$\text{CH}_3\text{F}$	302.16	4.65	296.24	3.31	293.5	2.8
$\text{CH}_2\text{F}_2$	306.96	9.45	299.76	6.81	296.3	5.6
$\text{CHF}_3$	311.92	14.41	303.27	10.34	299.0	8.3
$\text{CF}_4$	317.01	19.50	306.62	13.69	301.7	11.0

<sup>a</sup> Ref. [11].

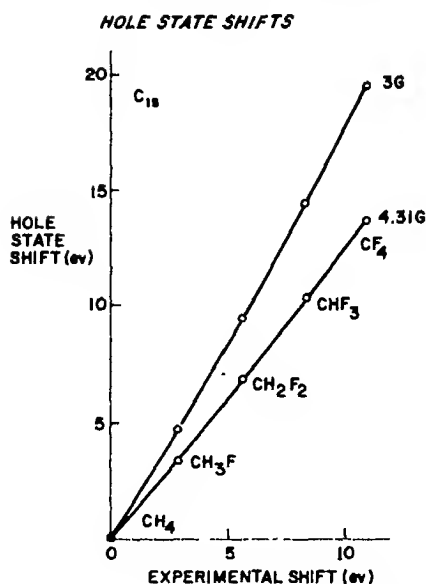
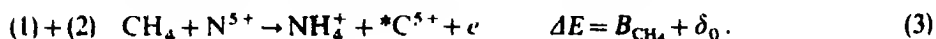


Fig. 2 Plot of hole state shifts (w.r.t.  $\text{CH}_4$ ) versus experimental shifts for the fluoromethanes as a function of basis set

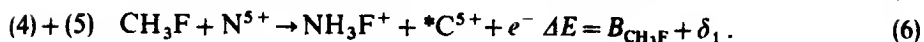
obtained from the difference in energy between the molecule and the core hole states are listed in Table 3 and the shifts are illustrated in Fig. 2. These calculations take into account electronic reorganization on core ionization and give binding energies which are in better agreement with the experimental values than are the Koopmans' theorem energies. However, for the 3 G and 4.31 G calculations the prediction of the shifts is not as good as the Koopmans' prediction, but for a double zeta calculation on the ground states and core hole states of  $\text{CH}_{4-n}\text{F}_n$  ( $n=0$  to 3) Brundle, Robin and Basch [19] have shown that the shifts are predicted with about equal accuracy by both methods.

### Equivalent Cores Shifts

The equivalent cores method of predicting shifts in core electron binding energies from thermodynamic data was developed by Jolly and Hendrickson [5]. Where thermodynamic data are available this method gives good predictions of shifts [20]. The principle of equivalent cores may be stated: "When a core electron is removed from an atom molecule or ion, the valence electrons adjust as if the nuclear charge of the atom had increased by one unit". Consider, for example, the shift in  $\text{C}_{1s}$  binding energy between  $\text{CH}_4$  and  $\text{CH}_3\text{F}$ .



Similarly for  $\text{CH}_3\text{F}$



The shift in core binding energies is thus given by



It is assumed in the equivalent cores approximation that the values  $\delta_0$  and  $\delta_1$  are small since the species  ${}^*\text{CH}_4^+$  and  $\text{NH}_4^+$ ,  ${}^*\text{CH}_3\text{F}^+$  and  $\text{NH}_3\text{F}^+$ , and  ${}^*\text{C}^{5+}$  and  $\text{N}^{5+}$  are taken to be chemically equivalent. The shift in binding energy is therefore given by the heat of reaction for (7). However, this expression is still valid even if  $\delta_1$  and  $\delta_0$  are not zero provided that  $\delta_1 - \delta_0 = 0$ , i.e. provided that the heats of reaction for core exchange are independent of the molecular environment for a particular pair of elements.

The energies for reaction of the type (7) may be calculated from SCF calculations on the individual molecules in their ground states. Semi empirical calculations of the equivalent cores shifts using all valence electron SCF MO calculations have been qualitatively successful [20–22] and *ab initio* minimal Slater basis set calculations have been more successful [6, 9]. It is therefore of interest to determine how much improvement is obtained when an extended basis set is used. Since photoionization is rapid compared with nuclear motion the geometries for the nitrogen cations were taken to be identical with those of the isoelectronic molecules. (This is also convenient on computational grounds since many of the two electron integrals can be retained.)

The total energies required to calculate the equivalent cores shifts in  $\text{C}_{1s}$  binding energies are shown in Table 2 and the equivalent cores shifts are shown in Table 4 and Fig. 3. The shifts are predicted well even by the smaller basis set calculations and in this respect they show less dependence on the choice of basis set than do the Koopmans' theorem and hole state calculations. For the large basis set calculations the equivalent cores shifts and Koopmans' theorem shifts are closely similar and this near equality is also observed between hole state shifts and Koopmans' theorem shifts when a double zeta basis set is used [19].

Table 4. Equivalent cores shifts (eV)

	3 G	4.31 G	Large	Experimental*
$\text{CH}_4$	0.0	0.0	0.0	0.0
$\text{CH}_3\text{F}$	2.95	3.05	2.82	2.8
$\text{CH}_2\text{F}_2$	6.34	6.31	5.99	5.6
$\text{CHF}_3$	10.16	9.66	9.31	8.3
$\text{CF}_4$	14.37	12.92	12.64	11.0
$\text{CH}_3\text{Cl}$			1.77	1.6
$\text{CH}_2\text{Cl}_2$			3.39	3.1

\* Refs. [10] and [11].

## EQUIVALENT CORE SHIFTS

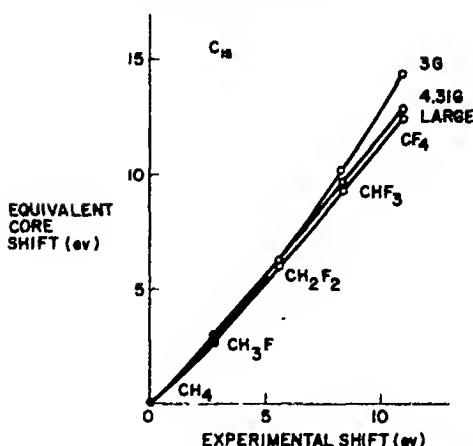
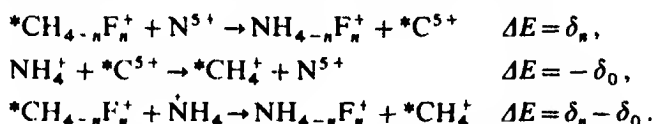


Fig. 3 Plot of equivalent core shifts (w.r.t.  $\text{CH}_4$ ) versus experimental shifts for the fluoromethanes as a function of basis set

It would therefore appear at least in the case of the fluoromethanes that all three methods of calculation tend towards the same results as the flexibility of the basis set increases.

The accuracy with which the equivalent cores method predicts shifts depends on how close the value of  $\delta_1 - \delta_0$  is to zero. It is therefore of interest to calculate the values of  $\delta_1 - \delta_0$  predicted by these calculations.



Values of  $\delta_n - \delta_0$  have been calculated from the 3 G and 4.31 G results (Table 2) and are shown in Table 5. Large deviations from zero occur with the 3 G calculations but the deviations in the case of the 4.31 G calculations are much smaller. It should be noted that the  $\delta_n - \delta_0$  values correspond to the difference in binding energy between the hole state and equivalent cores calculations and the deviation of  $\delta_n - \delta_0$  from zero actually acts as an improvement to the hole state calculations bringing them nearer to the experimental values. However, the deviation of  $\delta_n - \delta_0$  from zero is reduced greatly by the improvement of the basis set and a value of  $(\delta_n - \delta_0) = 0$  would mean that the equivalent cores and hole state calculations would predict the same shifts in core binding energies.

An experimental estimate of the value of  $\delta$  can readily be obtained for free atoms using ionization energy data to estimate the binding energy and comparing this with experimental measurements of the binding energy. Consider the following processes:

(1) 1s ionization in a carbon atom



Table 5. Calculated values of  $\delta_n - \delta_0$ 

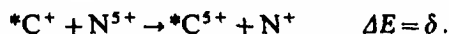
	n	$\delta_n - \delta_0$ (eV)	
		3 G	4.31 G
CH <sub>4</sub>	0	0.0	0.0
CH <sub>3</sub> F	1	-1.70	-0.26
CH <sub>2</sub> F <sub>2</sub>	2	-3.11	-0.49
CHF <sub>3</sub>	3	-4.24	-0.67
CF <sub>4</sub>	4	-5.13	-0.76

From calculations on  $^*C^{5+}$  and  $N^{5+}$  with corresponding basis sets, values of  $\delta_0$  may be obtained from the data given below

	Total energies (eV)	
	3 G	4.31 G
$^*C^{5+}$	- 483.6457	- 486.7477
$N^{5+}$	-1202.5878	- 1212.6748
$\delta_0$	- 13.78	- 9.24

Since in all cases only a single STO is used to describe the cores (being either a linear combination of three or four gaussians) it is clear that  $\delta_0$  is very sensitive to basis set. The detailed dependence on basis set will be discussed elsewhere.

## (2) Exchange of $^*C^{5+}$ core for equivalent $N^{5+}$ core



The sum of these reactions gives an estimate of the  $C_{1s}$  binding energy which will differ from the experimental value by  $\delta$



The energy of this reaction is the difference in the energy of the processes

$$C \rightarrow ^*C^{5+} + 5e^- \quad \Delta E_C = \sum_{i=1}^5 (IP_C)_i$$

$$N^+ \rightarrow N^{5+} + 4e^- \quad \Delta E_N = \sum_{i=2}^5 (IP_N)_i.$$

Hence

$$B_C + \delta = \Delta E_C - \Delta E_N = \sum_{i=1}^5 (IP)_C - \sum_{i=2}^5 (IP)_N$$

Using values of ionization potentials from Moores tables [23] gives

$$B_{C_{1s}} + \delta = 287.6 \text{ eV}$$

From gas phase measurements the  $1s$  binding energy of a carbon atom in benzene, i.e. one with approximately zero (CNDO) charge, is 290.4 eV [24]. This gives a value of  $\delta$  of about -2.8 eV. There will, however, be somewhat different electronic reorganization energies on ionization from a free atom and a

neutral atom in a molecule. Similar estimates of core binding energies and  $\delta$  have been made for a large number of atoms [25]. Shirley has recently made theoretical estimates of energies for core exchange in  $\text{CH}_4$  and  $\text{CH}_3\text{F}$ , and found the values of  $\delta = 6.0$  and  $5.9$  eV respectively [26]. The reason for this discrepancy between theory and experiment is uncertain.

### Chloromethanes

The Koopmans' theorem predictions, Table 1, overestimate the  $\text{C}_{1s}$  binding energy shifts and the error is larger than that obtained for  $\text{CH}_3\text{F}$  which has a  $\text{C}_{1s}$  shift, intermediate between  $\text{CH}_3\text{Cl}$  and  $\text{CH}_2\text{Cl}_2$ . The equivalent cores shifts are, however, in good agreement with the experimental values. We attribute the differing accuracy in the Koopmans' theorem results partly to the difficulty in obtaining fully compatible basis sets for the first and second row elements, and the improvement obtained from the equivalent cores calculations again illustrates the comparative lack of basis set dependency of this method.

## 4. Conclusion

These results, together with those of Brundle *et al.* [19] indicate that for the halomethanes when using large basis set calculations there is little difference between the shifts in core binding energies predicted by Koopmans' theorem, hole state and equivalent cores calculations. However if a minimal basis set is employed the best estimates of the shift are obtained from the equivalent cores calculations. The graphs (Figs. 1–3) illustrates clearly the order of decreasing basis set dependency of the predicted shifts to be:

Hole state > Koopmans' > Equivalent cores.

The results also suggest that for these closely related molecules differences in relaxation energies are small and therefore make only minor contributions to the shifts in binding energies. In this connection it is of interest to pursue the analysis of the relaxation (reorganization) energies as suggested by Gelius and Siegbahn [18]. The dominant contribution is that arising from the local charge distribution ( $E_A^{\text{contr}}$ ) and may be expressed as

$$E_A^{\text{contr}} = k'q_A + I'_A$$

where  $q_A$  is the charge on atom A before ionization,  $k'$  is a constant (2.5 eV in an atom [18, 27] and  $I'_A$  is the reorganization energy due to orbital contraction around a neutral atom in the molecule (13.7 eV for a carbon atom [18]). Estimates of the relaxation energy obtained from differences between binding energies calculated from Koopmans' theorem and hole states are shown in Table 6 together with atomic charges for the 4.31 G basis set calculations. These overall relaxation energies, which include  $E_A^{\text{low}}$ , are essentially constant. This is consistent with the tendency for Koopmans' theorem, hole state calculations and equivalent cores calculations to give the same estimates of shifts with flexible basis sets despite the fact that Koopmans' theorem neglects electronic relaxation while the hole states and equivalent cores calculations take it into account.

Table 6. Charges and relaxation energies (4.31 G basis set)

Molecule	Atom	Charge	Relaxation energy (eV) (Koopmans' B.E. - Hole state B.E.)
CH <sub>4</sub>	C	-0.875	11.4
	H	+0.219	-
CH <sub>3</sub> F	C	-0.281	11.2
	H	+0.226	-
CH <sub>2</sub> F <sub>2</sub>	F	-0.399	-
	C	+0.373	11.1
	H	+0.251	-
CHF <sub>3</sub>	F	-0.379	-
	C	+0.754	11.1
	H	+0.308	-
CF <sub>4</sub>	F	-0.354	-
	C	+1.328	11.3
	F	-0.332	-

It is unrealistic to compare directly the atomic relaxation energy data of Gelius and Siegbahn [18] with that calculated for the fluoromethanes because of differences in basis set. However, the prediction of a near constancy of relaxation energies for the fluoromethanes is interesting and from the analysis of Gelius and Siegbahn [18] this would only be expected if the sum of the charge dependent terms in  $E_A^{\text{contr}}$  and  $E_A^{\text{flow}}$  was constant. The charge dependent term in  $E_A^{\text{contr}}$  increases with increasing positive charge on the carbon atom along the series CH<sub>4</sub> to CF<sub>4</sub> and this implies that  $E_A^{\text{flow}}$  shows a similar charge dependency on the increasing total negative charge on the atoms bonded to carbon such that the sum total remains essentially constant.

*Acknowledgements.* Thanks are due to the Science Research Council for provision of a research studentship (D.B.A.) and to the Atlas Computer Laboratory for provision of computing facilities.

### References

1. Siegbahn, K., Nordling, C., Fahlman, A., Nordberg, R., Hamrin, K., Hedman, J., Johansson, G., Bergmark, T., Karlsson, S.-E., Lindgren, I., Lindberg, B.: ESCA-atomic, Molecular And Solid State Structure Studied By Means Of Electron Spectroscopy, Nova. Acta. Regiae Soc. Sci. Upsaliensis Ser. IV, Vol. 20, 1967
2. Siegbahn, K., Nordling, C., Johansson, G., Heden, J., Heden, P.F., Hamrin, K., Gelius, U., Bergmark, T., Werme, L. O., Manne, R., Baer, Y.: ESCA Applied To Free Molecules. North Holland 1969
3. Koopmans, T. A.: Physica 1, 104 (1933)
4. Bagus, P. S.: Phys. Rev. A **139**, 619 (1965)
5. Jolly, W. L., Hendrickson, D. N.: J. Am. Chem. Soc. **92**, 1863 (1970)
6. Clark, D. T., Adams, D. B.: J. Chem. Soc. Faraday Trans. II **68**, 1819 (1972)
7. Clementi, E., Popkie, H.: J. Am. Chem. Soc. **94**, 4057 (1972)
8. Cf. Aarons, L. J., Guest, M. F., Hillier, I. H.: J. Chem. Soc., Faraday Trans. II **68**, 1866 (1972) and references therein
9. Clark, D. T., Adams, D. B.: J. Electron Spectroscopy **1**, 302 (1972)
10. Thomas, T. D.: J. Am. Chem. Soc. **92**, 4184 (1970)
11. Davis, D. W., Shirley, D. A., Thomas, T. D.: J. Chem. Phys. **56**, 671 (1972)

12. Dunning, T.H.: *J. Chem. Phys.* **53**, 2823 (1970)
13. Veillard, A.: *Theoret. chim. Acta (Berl.)* **12**, 405 (1968)
14. Clementi, E.: *IBM Tech. Bull. San Jose*, 1970
15. Hillier, I.H., Saunders, V.R.: *Atlas Computer Laboratory*, 1972
16. Stewart, R.F.: *J. Chem. Phys.* **50**, 2485 (1969)
17. Clementi, E.: *IBM J. Res. Develop. Suppt.* **9**, 2 (1965)
18. Gelius, U., Siegbahn, K.: *Faraday Discuss. Chem. Soc.* **54**, 257 (1972)
19. Brundle, C. R., Robin, M., Basch, H.: *J. Chem. Phys.* **53**, 2196 (1970)
20. Jolly, W. L.: *Electron spectroscopy*, Shirley, D. A. (Ed.), p. 629. North Holland 1972 and references therein
21. Clark, D. T., Adams, D. B.: *Nature* **234**, 95 (1971)
22. Frost, D. C., Herring, F. G., McDowell, C. A., Woolsey, I. S.: *Chem. Phys. Letters* **13**, 391 (1972)
23. Moore, C. E.: *Atomic energy levels*. Washington: Nat. Bureau of Standards, Circ. 467 1949
24. Thomas, T. D.: *J. Chem. Phys.* **52**, 1373 (1970)
25. Adams, D. B., Clark, D. T.: In preparation
26. Shirley, D. A.: *Chem. Phys. Letters* **15**, 325 (1972)
27. Snyder, L. C.: *J. Chem. Phys.* **53**, 95 (1971)

Dr. D. T. Clark  
Department of Chemistry  
University of Durham  
Science Laboratories  
South Road  
Durham DH1 3LE, England



## Commentationes

# The Numerical Integration of a Molecular Integral Using Two Different Techniques\*

D. Rees, Deborah J. Moore, and P. R. Taylor

Mathematics Department, University of Nottingham

Received March 30, 1973

A two-electron integral which commonly occurs in molecular calculations is evaluated numerically using the different methods of Boys and Conroy and the results are discussed.

**Key words:** Numerical Integration

## 1. Introduction

The purpose of the investigation was to consider an integral with a singular integrand which occurs quite often in electronic calculations and to evaluate it using the "diophantine-type" methods of Boys *et al.* [1, 2] and Conroy [3]. The integral chosen is the two-electron, one-centre integral

$$\int \frac{e^{-\alpha r_1 - \beta r_2}}{r_{12}} dr_1 dr_2. \quad (1)$$

This integral has the advantage of being evaluated analytically to give

$$\frac{32\pi^2(\alpha^2 + 3\alpha\beta + \beta^2)}{\alpha^2\beta^2(\alpha + \beta)^3}, \quad (2)$$

thus making comparison with its numerical estimates possible.

## 2. The Co-Ordinate System

In both the numerical methods considered the spherical polar transformation of the co-ordinates of electron  $i$  ( $i = 1, 2$ ) suggested by Boys and Handy [2] was employed i.e.

$$\begin{aligned} r_i &= A_i(q_{r_i}/1 - q_{r_i}) \\ \Theta_i &= \pi(6q_{\Theta_i}^5 - 15q_{\Theta_i}^4 + 10q_{\Theta_i}^3) \\ \phi_i &= 2\pi q_{\phi_i}. \end{aligned} \quad (3)$$

This reduces the integral to the standard form

$$\int_0^1 \cdots \int_0^1 f(q_{r_1}, q_{\Theta_1}, q_{\phi_1}, q_{r_2}, q_{\Theta_2}, q_{\phi_2}) dq. \quad (4)$$

\* This paper was presented during the session on numerical integration methods for molecules of the 1970 Quantum Theory Conference in Nottingham. It has been revised in the light of the interesting discussion which followed.

Also in both the procedures the scale factors  $A_i$  ( $i = 1, 2$ ) were set equal to  $A$  in order to reduce the number of parameters.

### 3. Boys' Method

The transformation given by Eq. (3) transforms an integrand having a continuous derivative in polar co-ordinate space to an approximately periodic function. The Boys procedure gives an estimate of the integral by evaluating partial sums in which the integrand is evaluated at "diophantine" points in 3-space. Assuming that this procedure may be applied to integrands involving a singularity and generalizing it to the two-electron case, the partial sums are then given by

$$\frac{1}{N^2} \sum_{L_1=1}^N \sum_{L_2=1}^N f(L_1, L_2) \omega(L_1) \omega(L_2), \quad (5)$$

where

$$(q_r, q_r, q_r) = \left( \frac{L_i}{N}, \frac{DL_i}{N}, \frac{EL_i}{N} \right), \quad (i = 1, 2). \quad (6)$$

In (5),  $f$  is the integrand and  $\omega$  a one-electron weight which is obtained from the one-electron contribution to the Jacobian of the transformation. Equation (6) involves  $N$ , the number of integration points per electron, and also fixed constants  $D$  and  $E$  both of which depend on  $N$ .

### 4. The Singularity

In order to prevent infinite contributions which arise when  $r_{12} = 0$ , the method introduced by Boys and Cook [4] of coping with the singularity was used so that  $1/r_{12}$  is replaced by

$$\left( r_{12}^3 + \frac{\omega^4(L_1) \omega^4(L_2)}{Z} \right)^{-1/4}. \quad (7)$$

The only modification employed in (7) is to regard  $Z$  as a parameter rather than setting  $Z = 3$  as suggested by Boys and Cook.

### 5. Results

The numerical estimate of the integral using Boys' method depends therefore on two parameters  $A$  and  $Z$ .

Figure 1 shows the variation with  $Z$  of the percentage errors in the estimates of the integral for  $\alpha = 1$ ,  $\beta = 2$ ; in this case the scale factor was fixed at 2 (c.f. Fig. 2). For the range  $0 < Z < 25$ , the integral estimates vary quite markedly with  $Z$  and give the correct result when  $Z \simeq 16$ . It would therefore seem that the estimate of  $Z$  as approximately 3 arrived at by Boys and Cook using an electrostatic argument is not as satisfactory as  $Z = 16$  as a parameter in the approximation for the  $1/r_{12}$  singularity. In  $Z > 16$  the estimates become less accurate. This is to be expected since (7) approximates very closely to  $1/r_{12}$  for large  $Z$  and thus introduces further inaccuracies into the calculation.

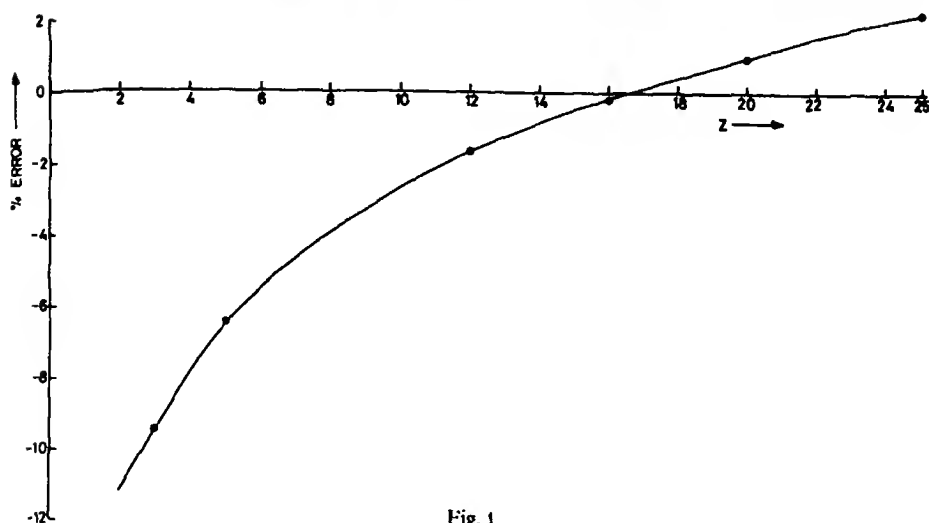


Fig. 1

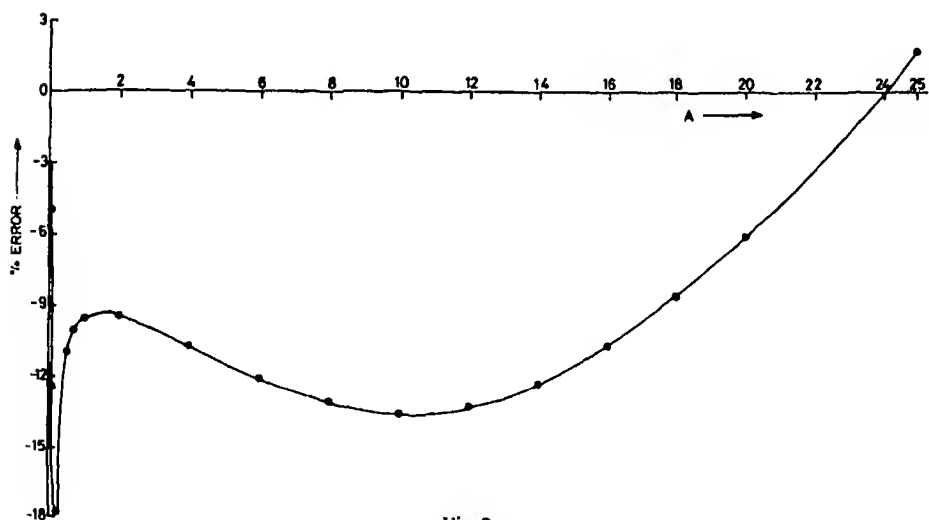


Fig. 2

Figure 2 shows the variation of the percentage error of the estimate with scale factor  $A$ . Again  $\alpha = 1$ ,  $\beta = 2$  and  $Z$  was chosen as the Boys and Cook estimate of 3. The error varies markedly with  $A$ , and the best estimates have an error of about 9-10% in the region  $1 \leq A \leq 2$ . It is in this region that the integral estimates are least sensitive to the scale factor.

Further calculation shows that if  $Z$  is increased to 16 the errors become much smaller and the integral estimates become much less sensitive to  $A$ . For example, when  $Z = 16$ , the percentage error is about 0.05 for  $1 < A < 6$ . These results were computed for  $N = 80$  - i.e. using 6400 grid points. The previous results were

repeated for several different values of the parameters ( $\alpha, \beta$ ) confirming the fact that the choice of  $Z = 3$  was far too small and that  $Z$ , in the neighbourhood of 16, gave much more accurate values of the estimates which were relatively insensitive to the choice of scale factor.

### 6. Conroy's Method

The method of Conroy [2] is based essentially on the approximation to this integral by the partial sum

$$\frac{1}{M} \sum_{L=1}^{M-1} g(L\alpha - [L\alpha]), \quad (8)$$

where  $g$  is the product of the integrand and the Jacobian of the transformation;  $M$  is the number of sample points. In expression (8)

$$\alpha = p/M, \quad (9)$$

where  $p$  is a six dimensional vector which is selected, having first chosen  $M$ , according to Conroy's prescription in which he lists optimised sets of  $M$  each with their respective optimised vector  $p$ . In (8) because the integral parts are subtracted off each component of  $L\alpha$ , the grid points lie inside the unit hypercube as is required from the transformation given by (3) (see also expression (4)).

The function  $g$  in expression (8) is evaluated in a six dimensional configuration space as opposed to the two three dimensional spaces of each electron as in the Boys method. It is therefore possible to order the components of  $p$  to correspond to the components of the vector  $q$  given by (3), in several different ways. Corresponding respectively to each of the components  $p_1, p_2, \dots, p_6$ , the ordering  $q_{r_1}, q_{\theta_1}, q_{\phi_1}, q_{r_2}, q_{\theta_2}, q_{\phi_2}$  and also the ordering  $q_{r_1}, q_{r_2}, q_{\theta_1}, q_{\theta_2}, q_{\phi_1}, q_{\phi_2}$  were considered. It was generally found, however, that the former ordering gave better results than the latter, despite the fact that the latter ordering allows both the radial co-ordinates of the electrons to assume greater importance in the integral. This first ordering has been used in all subsequent results.

One further point to note is that this integration procedure is such that the choice of integration points does not make  $r_{12}$  small so that no special device for coping with the singularity need be introduced. The exceptional case, however, is that for which  $M$  is even in which case the point corresponding to  $L = M/2$  makes  $r_{12}$  zero. In this case therefore the integration point corresponding to  $L = M/2$  is omitted.

### 7. Results

Figure 3 shows the variation with the scale factor  $A$  of the percentage error of the integral estimate for a choice of 6044 and 9644 grid point respectively. For the choice of 6044 points the results are relatively insensitive to the scale factor and are particularly good for  $2.5 \leq A \leq 10$  where the error is less than 1%. However, if the number of points used is increased to 9644 the results are much more sensitive to the scale factor particularly at values of  $A$  greater than 10 where there is marked oscillation. In this case the optimum results again occur in the region  $2.5 \leq A \leq 10$  but, despite the increased number of points, there is no improvement in the accuracy.

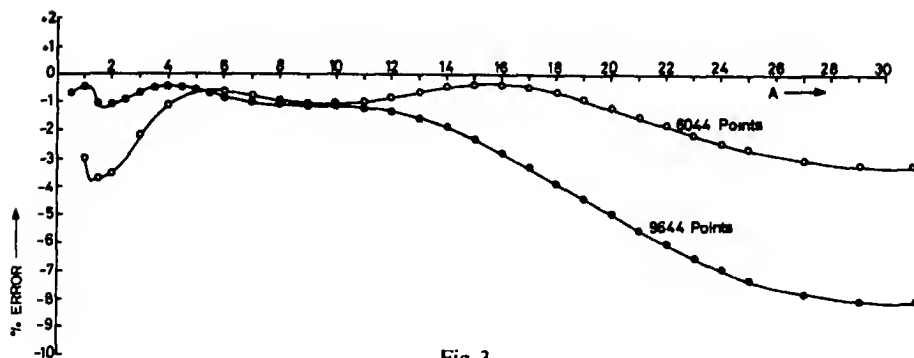


Fig. 3

### 8. Discussion and Conclusion

Both the methods of Conroy and Boys show that the results obtained are sensitive to the choice of scale factor (even for about 6000 grid points). It is, however, possible in estimating this integral to find for both methods ranges of  $A$  in which the estimates are optimum and least sensitive to the scale factor. In order to investigate the use of a different transformation from that given in (3), the alternative transformation

$$r_i = A(q_i/1 - q_i)^2 \quad (i = 1, 2) \quad (10)$$

was studied. The quadratic dependence of the  $r_i$  on  $q_i$  did in fact produce results which were less sensitive to  $A$  in both methods but the estimates were not as accurate. This alternative transformation did also predict the optimum value of  $Z$  for the Boys method as being approximately 22. Again confirming the choice of  $Z = 3$  as being far too low. Concerning the Boys method generally, it is accepted that arguments in favour of this method as applied to a six dimensional integral are not as strong as those favouring its application to one of three dimensions, particularly if  $r_{12}$  occurs explicitly, and as a singularity, making the integrand and its derivatives no longer continuous. However it does seem that, although the method was not originally conceived with the idea of application to integrals with singularities such as  $1/r_{12}$ , this suggestion of coping with the singularity can be improved by using optimised parameters.

Whereas the Boys method is essentially a one-electron, or three dimensional procedure, Conroy's method is essentially a configuration-space procedure. For 6044 grid points, Conroy's procedure gave exceedingly accurate results for certain ranges of the scale factor. Further it has the advantage of requiring no special devices to deal with the singularity. However, the fact that at 9644 grid points the variation of the estimates with  $A$  were much larger than that for 6044 raises some quite serious objections. Any desirable numerical integration method using 9000 points should produce results which are relatively insensitive to the scale factor  $A$ . The integral for the Boys procedure was not estimated using 9000 points so that no comparison can be made.

In conclusion, *ab initio* molecular calculations in which very many integrals are numerically computed will certainly need procedures requiring a reduced number of integration points if the computing time is to be minimized.

The possibility of using optimum parameters necessitating a smaller number of grid points in numerical integration techniques should therefore be further explored.

*Acknowledgments.* The authors would like to thank Mr. Brian Ford and Mr. C. E. Solomon for their valuable help in the writing and execution of the programs.

### References

1. Boys, S. F., Rajagopal, P.: *Adv. Quantum Chem.* **2**, 1 (1965)
2. Boys, S. F., Handy, N. C.: *Proc. Roy. Soc. (London)* **A311**, 309 (1969)
3. Conroy, H.: *J. Chem. Phys.* **47**, 5307 (1967)
4. Boys, S. F., Cook, G. B.: *Rev. Mod. Phys.* **32**, 285 (1960)

Dr. D. Rees  
Mathematics Department  
University of Nottingham  
Nottingham, England

# A Six-Dimensional Quadrature Procedure\*

J. Hyslop

Department of Mathematics  
University of Technology, Loughborough

Received March 30, 1973

An account is given of the use of Gaussian quadrature product formulae in the evaluation of certain six-dimensional, two-centre integrals involving one-electron Green's functions. These integrals occur in a new molecular variational principle recently proposed by Hall, Hyslop and Rees [1] from which an approximate energy may be derived which can be shown to be at least as good as that obtained from the Rayleigh-Ritz principle. Reductions in computing time are realized by removing certain singularities using a subtraction technique and also by using an empirically determined Richardson-type extrapolation formula.

**Key words:** Gaussian quadrature – Green's functions – Singularities

## 1. Introduction

In a recent paper Hall, Hyslop and Rees [1] proposed a variation principle for molecular energies and applied it to the calculation of an upper bound for the ground state energy of the  $H_2^+$  molecule. The six-dimensional Green's function integrals involved were evaluated by employing a semi-analytical technique in which a Fourier transform representation of the Green's function enabled the integrals to be reduced to a single quadrature in the simplest cases considered and to a triple quadrature in more complicated cases. However, generalization of this method is not straightforward and indeed as pointed out in [1], some of the additional integrals required in an alternative functional cannot be evaluated by the techniques employed on the original functional.

The functional contained in the original principle was derived using the usual Born-Oppenheimer approximation that the nuclear and electronic coordinates could be completely separated, whereas the alternative functional proposed did not assume this separation. Since a direct comparison of these functionals is of interest and also since an integration routine which is more readily generalized is necessary the possibility of direct numerical evaluation of the integrals is considered.

## 2. Formulation of the Integrals

On introducing a class  $\psi(x)$  of scaled trial functions, where the scaled variables are denoted by  $x$  with  $x = kr$  when the electronic energy is given by

\*This paper was presented during the session on numerical integration methods for molecules of the 1970 Quantum Theory Conference in Nottingham. It has been revised in the light of the interesting discussion which followed.

$E = -k^2/2$ , the functional used in [1] may be written as

$$k = \iint \psi^*(x_1) V(x_1) G(x_1, x_2) V(x_2) \psi(x_2) dx_1 dx_2 / \int \psi^*(x) V(x) \psi(x) dx \quad (1)$$

$G(x_1, x_2) = -\exp(-x_{12})/(2\pi x_{12})$  is the scaled Green's function with  $x_{12} = |x_1 - x_2|$  and  $V(x)$  is the scaled electronic potential energy corresponding to

$$V(r) = -(1/r_a + 1/r_b) \quad (2)$$

where  $r_a$  and  $r_b$  are the distances between the electron and protons  $a$  and  $b$ , whose separation is  $R$ . The total energy is then given by

$$W = -k^2/2 + 1/R \quad (3)$$

The alternative functional may be written as

$$k' = \iint \psi^*(x_1) V'(x_1) G(x_1, x_2) V'(x_2) \psi(x_2) dx_1 dx_2 / \int \psi^*(x) V'(x) \psi(x) dx \quad (4)$$

in which  $V'(x)$  is the scaled total potential energy corresponding to

$$V'(r) = -(1/r_a + 1/r_b) + 1/R \quad (5)$$

and the total energy corresponding to this functional is

$$W' = -k'^2/2 \quad (6)$$

The values of  $k$  are then obtained as functions of the independent variable  $\varrho = kR$  from equation (1) with corresponding internuclear separation  $\varrho/k$ . Correspondingly  $k'$  is given as a function of  $\varrho' = k'R$  by equation (4) at the internuclear separation  $\varrho'/k'$  and hence energy curves may be obtained and compared.

Two centre elliptic coordinates  $(\lambda, \mu, \phi)$  are used where

$$\lambda = (r_a + r_b)/R, \quad \mu = (r_a - r_b)/R \quad (7)$$

and  $\phi$  is the azimuthal angle about the internuclear axis. By way of illustration the trial function used is the simple united atom approximation

$$\psi(x) = \exp(-c\lambda) \quad (8)$$

where  $c$  is a variable parameter. This function may be regarded as being already scaled since it involves only the ratio of distances.

On writing the functionals  $k$  and  $k'$  as  $I/J$  and  $I'/J'$  respectively it is easily seen that the normalization integrals  $J$  and  $J'$  are given by

$$J(\varrho, c) = -\frac{1}{2}\pi (\varrho/c)^2 (1 + 2c) \exp(-2c) \quad (9a)$$

and

$$J'(\varrho', c) = -\frac{1}{24}\pi (\varrho^2/c^3) (20c^2 + 6c - 3) \exp(-2c) \quad (9b)$$

and that the main integrals  $I$  and  $I'$  may be written as

$$I = -\frac{1}{4}\varrho^4 \int x_{12}^{-1} \exp(-x_{12}) \exp[-c(\lambda_1 + \lambda_2)] \lambda_1 \lambda_2 d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\phi \quad (10a)$$

and

$$I' = -\frac{1}{4}\varrho'^4 \int x_{12}^{-1} \exp(-x_{12}) \exp[-c(\lambda_1 + \lambda_2)] [\lambda_1 - (\lambda_1^2 - \mu_1^2)/4] \cdot [\lambda_2 - (\lambda_2^2 - \mu_2^2)/4] d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\phi \quad (10b)$$



The scaled distance  $x_{12}$  is given by

$$4x_{12}^2/\varrho^2 = (\lambda_1^2 - 1)(1 - \mu_1^2) + (\lambda_2^2 - 1)(1 - \mu_2^2) + (\lambda_1\mu_1 - \lambda_2\mu_2)^2 - 2[(\lambda_1^2 - 1)(1 - \mu_1^2)(\lambda_2^2 - 1)(1 - \mu_2^2)]^{1/2} \cos \phi \quad (11)$$

where  $\phi = \phi_1 - \phi_2$  is the difference in azimuthal angles.

In the  $I'$  integral, therefore,  $\varrho$  is replaced by  $\varrho'$  and the  $\lambda_1\lambda_2$  term in  $I$  is replaced by

$$[\lambda_1 - (\lambda_1^2 - \mu_1^2)/4][\lambda_2 - (\lambda_2^2 - \mu_2^2)/4] \quad (12)$$

### 3. Numerical Integration Procedure

The formula used for the integrations is a five-dimensional Gaussian product formula of the type given by Stroud and Secrest [2]. For the  $\mu$  integrations the basic result is

$$\int_{-1}^{+1} f(\mu) d\mu = \sum_{i=1}^N A_\mu(i) f[x_\mu(i)] \quad (13)$$

where  $x_\mu(i)$  is the  $i^{\text{th}}$  grid-point and  $A_\mu(i)$  the corresponding weight for  $N$ -point Gauss-Legendre quadrature. In the case of the  $\lambda$  integrations the result is

$$\int_1^\infty f(\lambda) \exp(-c\lambda) d\lambda = c^{-1} \exp(-c) \sum_{i=1}^N A_L(i) f[1 + x_L(i)/c] \quad (14)$$

in which  $x_L(i)$  and  $A_L(i)$  are the grid-points and weights for  $N$ -point Gauss-Laguerre quadrature. The quantities  $x_\mu(i)$ ,  $A_\mu(i)$ ,  $x_L(i)$  and  $A_L(i)$  are extensively tabulated by Stroud and Secrest. The azimuthal integration is carried out by means of the Chebyshev-type result

$$\int_0^{2\pi} f(\cos \phi) d\phi = \frac{2\pi}{N} \sum_{i=1}^N f\left\{\cos\left[\frac{(2i-1)\pi}{2N}\right]\right\} \quad (15)$$

which is also referred to by Kopal [3] as the  $N$ -point Gauss-Mehler quadrature formula.

The main difficulty associated with the use of Gaussian quadrature formulae in the evaluation of molecular integrals of the type  $I$  or  $I'$  is the extremely slow convergence with increasing  $N$  because of the  $x_{12}^{-1}$  singularity. Extensive numerical tests were carried out using the technique of Boys and Rajagopal [3] for the removal of the singularity, but, even when various modifications were used, the convergence was still found to be exceedingly slow. It is proposed therefore to consider numerically the integrals  $S$  and  $S'$  rather than  $I$  and  $I'$  where

$$S = \frac{1}{4}\varrho^4 \int x_{12}^{-1} [1 - \exp(-x_{12})] \exp[-c(\lambda_1 + \lambda_2)] \lambda_1 \lambda_2 d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\phi \quad (16)$$

with a similar expression for  $S'$  with  $\varrho$  replaced by  $\varrho'$  and  $\lambda_1\lambda_2$  replaced by (12). The singularity of  $x_{12}=0$  is now removable and numerical tests confirm that a considerable improvement in convergence is realized when an  $N^5$ -point Gaussian formula is used.

It is necessary then to evaluate the complementary integral

$$A = \frac{1}{4} q^4 \int x_{12}^{-1} \exp[-c(\lambda_1 + \lambda_2)] \lambda_1 \lambda_2 d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\phi \quad (17)$$

with a corresponding expression for  $A'$ , using (12). These integrals may be treated analytically using the Neumann expression of  $x_{12}^{-1}$  in terms of  $(\lambda, \mu, \phi)$  as quoted, for example, by Harris and Michels [4]. The analysis is similar to that used by Sugiura [5] in his classical evaluation of the Heitler-London hydrogen molecule integrals and analytical expressions for  $A$  and  $A'$  are readily obtained in terms of the subsidiary integrals

$$L_n(c) = \int_1^\infty \lambda^n \exp(-c\lambda) \log(\lambda + 1)/(\lambda - 1) d\lambda, \quad (18)$$

$n$  ranging from 0 to 4. The values of  $L_n(c)$  are obtained by successive differentiations of the result

$$L_0(c) = e^{-c} [(\gamma + \log 2c) \exp(-c) + \exp(c) E_1(2c)] \quad (19)$$

where  $\gamma = 0.577215\dots$  is Euler's constant and  $E_1(z)$  is the exponential integral

$$E_1(z) = \int_1^\infty t^{-1} \exp(-zt) dt \quad (20)$$

as defined by Abramowitz and Stegun [6].

The integrals  $I$  and  $I'$  are then given by subtraction to be  $I = S - A$  and  $I' = S' - A'$  (21)

once  $S$  and  $S'$  have been evaluated numerically and  $A$  and  $A'$  evaluated analytically.

#### 4. Improvement in Convergence

Denoting the value of  $S$  obtained from the  $N^3$ -point Gauss formula by  $S_N$ , numerical tests indicate that the convergence of the required  $\{S_N\}$  is rapidly approaching geometric, so that the Aitken  $\delta^2$ -extrapolation procedure may be used to speed up the process, Noble [7]. In fact, using a combination of the Richardson and Aitken extrapolation techniques as outlined in [7] it appears that a formula of the form

$$S \simeq S_{N+1} - a(S_N - S_{N+1}) \quad (22)$$

is valid for large  $N$ , certainly for  $N$  greater than 8.

In order to reduce the value of  $N$  an empirical formula of the form

$$S \simeq b[S_5 - a(S_4 - S_5)] \quad (23)$$

was investigated and extensive numerical tests indicated that an optimum value for the constant  $a$  could be found which was such that the correction factor  $b$  was extremely slowly varying with  $q$  and the corresponding value of  $c$  in the region of its optimum value, as prescribed by section 5. Some representative values demonstrating the validity of this formula are shown in the following table for both  $S$  and  $S'$ .

Table 1. Variation of the correction factor  $b$  in the extrapolation formula

$\rho$	$c$	$S$ integrals $a = 0.8$	$S'$ integrals $a = 0.9$
1.0	0.48	0.99843	0.99777
2.0	0.93	0.99876	0.99891
3.0	1.36	0.99882	0.99902
4.0	1.80	0.99880	0.99902
5.0	2.23	0.99869	0.99896
6.0	2.67	0.99854	0.99887

Further numerical tests on extrapolation formula (23) indicate that both  $S$  and  $S'$  may be obtained to an accuracy of better than 3 parts in  $10^4$  for values of  $\rho$  (or  $\rho'$ ) greater than 1.0 with a maximum value of  $N = 5$ , that is, 3125 Gauss-points for the five-dimensional integrals. For  $\rho \leq 1.0$ , it is necessary to increase  $N$ , but in this case since the united atom limit is being rapidly approached, the resulting electronic energy values are very close to the exact results, of Dalgarno and Poots [8]. Indeed, in this region the behaviour of the total potential energy curve is dominated by the  $1/R$  nuclear repulsion term.

### 5. Optimization Routine

The technique is similar to that used in [1]. Thus, for a given  $\rho$  the functional  $k$  is optimized with respect to the parameter  $c$  according to

$$(\partial k / \partial c)_\rho = 0 \quad (24)$$

using a simple one-dimensional search routine based on extrapolation formula (23) coupled with analytical evaluation of  $A$  and  $A'$ . Various checks on accuracy are carried out involving increases in the number of Gauss points. This yields optimum values  $k_{\text{opt}}$  and  $c_{\text{opt}}$  with corresponding  $R$  values given by  $\rho/k_{\text{opt}}$  and  $W$  given by

$$W_{\text{opt}}(\rho) = -k_{\text{opt}}^2/2 + 1/R \quad (25)$$

Similarly the functional  $k'$  is optimized according to

$$(\partial k' / \partial c)_{\rho'} = 0 \quad (26)$$

and yields  $k'_{\text{opt}}$ ,  $c_{\text{opt}}$  and correspondingly  $R' = \rho'/k'_{\text{opt}}$  and

$$W'_{\text{opt}}(R') = -k_{\text{opt}}'^2/2 \quad (27)$$

The equilibrium separation  $R_0$  is obtained by optimizing  $W(\rho, c)$  or  $W'(\rho', c)$  as in [1] using the direct search routine of Rosenbrock [9], making use once again of the empirical extrapolation formula (23) to reduce computing time.

### 6. Discussion

The main limitation on accuracy is the numerical evaluation of the six-dimensional integrals  $S$  and  $S'$ . The use of the empirically determined extrapolation formula enables a maximum error of  $3 \times 10^{-4}$  to be realized with a

total of 3125 Gauss-points in the determination of these integrals. Cancellation effects in evaluating  $(S - A)$  and  $(S' - A')$  and also between  $-k^2/2$  and  $1/R$  may increase the error to a maximum of  $3 \times 10^{-3}$  in the total energies. The energy curves  $E(R)$  and  $E'(R')$  and  $W(R)$  and  $W'(R')$  are found to be indistinguishable to this accuracy and it is concluded that the Born-Oppenheimer separation in the functionals is valid within the limits of the accuracy of the present calculations. As an illustration, the optimum values of  $W = -0.590$  at  $R = 1.912$  as compared with  $W' = -0.593$  at  $R' = 1.909$  may be quoted and compared with the exact result of  $-0.603$  at  $2.00$ .

### References

1. Hall, G. G., Hyslop, J., Rees, D.: *Intern. J. Quantum Chem.* **4**, 5 (1970)
2. Stroud, A. H., Secrest, D.: *Gaussian quadrature formulas*. Englewood Cliffs, New Jersey: Prentice-Hall, 1966
3. Boys, S. F., Rajagopal, P.: *Adv. Quantum Chem.* **2**, 1 (1965)
4. Harris, F. E., Michels, H. H.: *Adv. Chem. Phys.* **13**, 205 (1967)
5. Sugiura, Y.: *Z. Physik* **45**, 484 (1927)
6. Abramowitz, M., Stegun, I. A.: *Handbook of mathematical functions*. New York: Dover Publications (1965)
7. Noble, B.: *Numerical methods*. Edinburgh: Oliver and Boyd (1964)
8. Dalgarno, A., Poots, G.: *Proc. Phys. Soc. (London)* **A 67**, 343 (1954)
9. Rosenbrock, H. H.: *Computer Journal* **3**, 175 (1960)

Dr. J. Hyslop  
Department of Mathematics  
University of Technology  
Loughborough, England

## Integration Points for the Reduction of Boundary Conditions\*

N. C. Handy and S. F. Boys

Theoretical Chemistry Department, University Chemical Laboratory  
Lensfield Road, Cambridge, CB2 1EW

Received March 30, 1973

An analysis of a method for the numerical evaluation of the integral  $\int_a^b f(x) dx$  is presented. The method introduces a change of variable,  $x = x(q)$ , with the property that  $d^n x/dq^n$  is zero at  $x = a$ ,  $x = b$  for  $n = 0, 1, 2, \dots, N$ , where  $N$  is an integer to be chosen. The Euler-Maclaurin formula shows that the resulting integral in the variable  $q$  is ideally suited for numerical integration, using equally spaced points and equal weights in  $q$ -space. Examples are given for various integrals which occur in quantum chemistry and applications to more than one dimension are discussed.

*Key words:* Numerical integration

For the numerical integration of many dimensional integrals a change of variable to make zero various low order derivatives at the boundaries has been variously used. [See Boys and Rajagopal (1965), Boys and Handy (1969) and a special scheme by Sag and Szekeres (1964), where a device which makes all boundary derivatives equal to zero is used. The latter is similar to the  $T_\infty$  transformation introduced below.] For an approximate estimate of the errors in transformations of such types, the most useful data appear to be the results for various one dimensional integrals. These do not appear to have been available and such a set is given here. Since these transformations are the simplest way of making various derivatives have zero values at the boundaries of the integration range, they will be referred to here as boundary derivative reductions.

The type of transformations with which we are concerned is the simplest type of change of variable  $x(q)$  with

$$\int_0^1 F(x) dx = \int_0^1 F(x(q)) \frac{dx}{dq} dq \equiv \int_0^1 G(q) dq. \quad (1)$$

It is always arranged that  $dx/dq$  varies as  $q^n$  near  $q = 0$  and  $(1 - q)^n$  near  $q = 1$ . Then the simplest numerical approximation to this with  $n + 1$  equally spaced

\* This paper was presented during the session on numerical integration methods for molecules of the 1970 Quantum Theory Conference in Nottingham. It has been revised in the light of the interesting discussion which followed.

points (including the two limits) is

$$\begin{aligned} & \sum_{i=1}^{n-1} G(q_i \equiv i/n) + 0.5(G(0) + G(1)) \\ &= \int_0^1 F(x) dx - \frac{n^{-2}}{12} (G^{(1)}(1) - G^{(1)}(0)) + \frac{n^{-4}}{720} (G^{(3)}(1) - G^{(3)}(0)) \\ & \quad - \frac{n^{-6}}{30240} (G^{(5)}(1) - G^{(5)}(0)) + O(n^{-8}). \end{aligned} \quad (2)$$

The terms varying as  $n^{-2}$ ,  $n^{-4}$  etc. are the correction terms given by the Euler-Maclaurin formula [see Krylov (1962) or Whittaker and Robinson (1937)]. But these depend on the values of the odd derivatives at the boundary. If  $a+b > 1$ , where  $q^b$  denotes the variation of  $F(x(q))$  for small  $q$ , then the  $n^{-2}$  term vanishes; if further  $a+b > 3$ , then the  $n^{-4}$  term vanishes and so on. In the transformations  $T_3$  and  $T_5$  used below, we generally have residuals of  $n^{-4}$  and  $n^{-6}$  respectively. The integrals without transformation generally give  $n^{-2}$  errors and these are given for comparison.

It is considered that for the evaluation of many dimensional integrals, the use of such transformations for each dimension separately is probably the most desirable first step for most integrals. In such a case it is desirable to be able to estimate the value of the errors to be expected. Such errors for  $n=6$  to about  $n=12$  in the tables provide an approximate means of doing this and seem to be the only record of suitable quantities for this.

The notations  $T_p$  with  $p$  giving the lowest power of  $q$  are used in the tables for the following transformations with  $T_1$  thus denoting no transformation:

$$\begin{aligned} T_1: & x = q \\ T_2: & x = a_2 \int_0^1 q(1-q) dx = 3q^2 - 2q^3 \\ T_p: & x = a_p \int_0^1 q^p (1-q)^{p-1} dq \\ T_\infty: & x = [\exp(-q^{-1})] [\exp(-q^{-1}) + \exp(-(1-q)^{-1})]^{-1}. \end{aligned} \quad (3)$$

The constants  $a_p$  are chosen to make  $x$  cover the range 0 to 1 as  $q$  covers the same range. The numerical integration for a given  $F(x)$  is then performed through Eq. (1) and through

$$\left( \int_0^1 G(q) dq \right)_{\text{approximate}} = \sum_{i=1}^{n-1} G(i/n) + 0.5(G(0) + G(1)). \quad (4)$$

The fractional errors  $\varepsilon$  obtained by these transformations are given for different  $n$ , all multiplied by  $10^5$ , so that 1. denotes a normal high accuracy. In each table a quantity (error  $\times (n^p)$ ) is recorded as  $( )_p$ . This is the value of  $C$  if the error varied as  $Cn^{-p}$ . Hence the constancy of  $( )_p$  in a column denotes a close approximation to an  $n^{-p}$  law. The fractional error results are stated to 0.00001 but there may be errors of about 0.00002 and so any case less than 0.00005 has been given as 0. The accurate standard of reference was taken from  $n=100$ ,  $T_\infty$ , and so this has no entry.

Table 1. Fractional errors  $\times 10^5$  for  $F = \exp(-4(x-0.5)^2)$ 

$n$	$T_1$	$T_3$	$T_5$	$T_\infty$
6	-915.6 (-0.3296) <sub>2</sub>	-38.46 (-0.4985) <sub>4</sub>	311.4 (145.3) <sub>6</sub>	-196.3
8	-514.2 (-0.3291) <sub>2</sub>	-13.59 (-0.5566) <sub>4</sub>	15.42 (40.42) <sub>6</sub>	-116.7
10	-328.8 (-0.3288) <sub>2</sub>	-5.368 (-0.5368) <sub>4</sub>	1.345 (13.45) <sub>6</sub>	-3.632
12	-228.3 (-0.3287) <sub>2</sub>	-2.526 (-0.5239) <sub>4</sub>	0.3356 (10.02) <sub>6</sub>	12.10
16	-128.3 (-0.3286) <sub>2</sub>	-0.7790 (-0.5106) <sub>4</sub>	0.05830 (9.781) <sub>6</sub>	0.5658
20	-82.13 (-0.3285) <sub>2</sub>	-0.3151 (-0.5042) <sub>4</sub>	0.01531 (9.799) <sub>6</sub>	-0.4428
40	-20.52 (-0.3284) <sub>2</sub>	-0.01938 (-0.4962) <sub>4</sub>	0.00021 (8.599) <sub>6</sub>	-0.00075
60	-9.122 (-0.3284) <sub>2</sub>	-0.00383 (-0.4965) <sub>4</sub>	0	0
100	-3.284 (-0.3284) <sub>2</sub>	-0.00049 (-0.4959) <sub>4</sub>	0	

Table 2. Fractional errors  $\times 10^5$  for  $F = \exp(x-0.5)^2$ 

$n$	$T_1$	$T_3$	$T_5$	$T_\infty$
6	543.6 (0.1957) <sub>2</sub>	-85.06 (-1.102) <sub>4</sub>	54.75 (25.54) <sub>6</sub>	-853.5
8	306.2 (0.1959) <sub>2</sub>	-27.74 (-1.136) <sub>4</sub>	8.781 (23.02) <sub>6</sub>	-271.7
10	196.1 (0.1961) <sub>2</sub>	-11.51 (-1.151) <sub>4</sub>	2.337 (23.37) <sub>6</sub>	0.9609
12	136.2 (0.1962) <sub>2</sub>	-5.589 (-1.159) <sub>4</sub>	0.7845 (23.42) <sub>6</sub>	28.65
16	76.66 (0.1962) <sub>2</sub>	-1.781 (-1.167) <sub>4</sub>	0.1399 (23.48) <sub>6</sub>	1.304
20	49.07 (0.1963) <sub>2</sub>	-0.7319 (-1.171) <sub>4</sub>	0.03671 (23.49) <sub>6</sub>	-1.046
40	12.27 (0.1963) <sub>2</sub>	-0.04597 (-1.177) <sub>4</sub>	0.00558 (22.85) <sub>6</sub>	-0.00176
60	5.454 (0.1963) <sub>2</sub>	-0.00909 (-1.179) <sub>4</sub>	0	0
100	1.963 (0.1963) <sub>2</sub>	-0.00117 (-1.177) <sub>4</sub>	0	

Table 3. Fractional errors  $\times 10^5$  for  $F = 0.5 \cosh(2x-1)$ 

$n$	$T_1$	$T_3$	$T_5$	$T_\infty$
6	942.2 (0.3327) <sub>2</sub>	-91.61 (-1.187) <sub>4</sub>	63.27 (29.51) <sub>6</sub>	1014.2
8	520.3 (0.3329) <sub>2</sub>	-30.33 (-1.242) <sub>4</sub>	9.834 (25.78) <sub>6</sub>	-301.3
10	333.1 (0.3331) <sub>2</sub>	-12.67 (-1.267) <sub>4</sub>	2.607 (26.07) <sub>6</sub>	2.296
12	231.4 (0.3332) <sub>2</sub>	-6.177 (-1.281) <sub>4</sub>	0.8752 (26.13) <sub>6</sub>	31.88
16	130.2 (0.3332) <sub>2</sub>	-1.975 (-1.294) <sub>4</sub>	0.1560 (26.18) <sub>6</sub>	1.447
20	83.31 (0.3333) <sub>2</sub>	-0.8133 (-1.301) <sub>4</sub>	0.04093 (26.20) <sub>6</sub>	-1.164
40	20.83 (0.3333) <sub>2</sub>	-0.05119 (-1.311) <sub>4</sub>	0.00625 (25.61) <sub>6</sub>	-0.00196
60	9.259 (0.3333) <sub>2</sub>	-0.01013 (-1.314) <sub>4</sub>	0	0
100	3.333 (0.3333) <sub>2</sub>	-0.00132 (-1.315) <sub>4</sub>	0	

The transformation  $T_\infty$  is a special form introduced because it has all the boundary derivatives zero. It can be inferred non-rigorously by inspection of the simple Euler-Maclaurin formula that the errors would decrease more rapidly than any finite power of  $n^{-1}$ . This can also be shown by more complicated vigorous mathematical argument. It will be seen that the results for  $T_\infty$  in every table appear to be in agreement with this.

The tables show how markedly improved are the results for  $T_3$  and  $T_5$  compared to  $T_1$  in all normal cases (Tables 1-3) and how closely they follow Euler-Maclaurin predictions. A few exceptional functions are included and again

Table 4. Fractional errors  $\times 10^5$  for  $F = \sin(\pi x)$ 

$n$	$T_1$	$T_2$	$T_3$	$T_\infty$
6	-2295.1 (-0.8262) <sub>2</sub>	112.6 (1.460) <sub>4</sub>	-21.31 (-9.942) <sub>6</sub>	405.5
8	-1288.4 (-0.8246) <sub>2</sub>	35.90 (1.471) <sub>4</sub>	-4.144 (-10.86) <sub>6</sub>	-10.41
10	-823.8 (-0.8238) <sub>2</sub>	14.74 (1.474) <sub>4</sub>	-1.118 (-11.18) <sub>6</sub>	-8.088
12	-571.8 (-0.8234) <sub>2</sub>	7.121 (1.476) <sub>4</sub>	-0.3803 (-11.35) <sub>6</sub>	0.2640
16	321.4 (-0.8230) <sub>2</sub>	2.256 (1.478) <sub>4</sub>	-0.06873 (-11.53) <sub>6</sub>	0.0456
20	-205.7 (-0.8228) <sub>2</sub>	0.9244 (1.479) <sub>4</sub>	-0.01815 (-11.62) <sub>6</sub>	-0.0104
40	51.40 (-0.8225) <sub>2</sub>	0.05779 (1.480) <sub>4</sub>	-0.00030 (-12.54) <sub>6</sub>	0
60	22.84 (-0.8225) <sub>2</sub>	0.01141 (1.478) <sub>4</sub>	0	0
100	8.224 (-0.8225) <sub>2</sub>	0.00148 (1.481) <sub>4</sub>	0	-

Table 5. Fractional errors  $\times 10^5$  for  $F = \exp(-100(x-0.5)^2)$ 

$n$	$T_1$	$T_3$	$T_5$	$T_\infty$
6	5727.5 (2.016) <sub>2</sub>	76371.1 (989.8) <sub>4</sub>	131406.8	88076.3
8	361.2 (2.312) <sub>2</sub>	33717.0 (1381.0) <sub>4</sub>	73641.6	41723.3
10	10.34 (0.1034) <sub>2</sub>	12746.7 (1274.6) <sub>4</sub>	39857.7	17265.9
12	0.1344 (0.0002) <sub>2</sub>	4038.0 (837.3) <sub>4</sub>	19857.7	6006.6
16	0	246.7 (161.6) <sub>4</sub>	3664.7	429.4
20	0	8.384 (13.41) <sub>4</sub>	464.9	15.5
40	0	0	0.00011	0
60	0	0	0	0
100	0	0	0	-

Table 6. Fractional errors  $\times 10^5$  for  $F = (1-x)^{-2} \exp(-(1-x)^{-1})$ 

$n$	$T_1$	$T_3$	$T_5$	$T_\infty$
6	-403.8 (-1.454) <sub>2</sub>	-706.5 (-9.156) <sub>4</sub>	-2813.0 (-1312.0) <sub>6</sub>	-2878.0
8	-242.2 (-1.550) <sub>2</sub>	232.4 (9.522) <sub>4</sub>	-404.2 (-1059.0) <sub>6</sub>	-760.1
10	-89.78 (-0.897) <sub>2</sub>	-71.83 (-7.183) <sub>4</sub>	288.4 (2884.0) <sub>6</sub>	-178.5
12	-47.36 (-0.682) <sub>2</sub>	15.00 (3.112) <sub>4</sub>	-64.12 (-1915.0) <sub>6</sub>	18.25
16	-31.83 (-0.815) <sub>2</sub>	0.2368 (1.552) <sub>4</sub>	4.178 (700.8) <sub>6</sub>	-9.824
20	21.24 (-0.850) <sub>2</sub>	-0.3258 (-5.257) <sub>4</sub>	0.9393 (601.1) <sub>6</sub>	-2.794
40	-5.209 (-0.833) <sub>2</sub>	-0.01958 (-4.968) <sub>4</sub>	0.00021 (8.583) <sub>6</sub>	-0.00142
60	-2.314 (-0.833) <sub>2</sub>	-0.00387 (-5.018) <sub>4</sub>	0	0
100	0	0	0	-

they are markedly in accordance with expectations. For example (Table 4) since  $\sin(\pi x)$  already behaves as  $x$  and  $1-x$  at the boundaries the transformations  $T_2$  and  $T_4$  carry it to the same state as  $T_3$  and  $T_5$  in the ordinary case.

In Table 5,  $\exp(-100(x-0.5)^2)$  is a very peaked function, and it is clear that the transformations each make it more peaked. Clearly the integral will not be accurate until the peak is accurately integrated.

In Table 7, the function  $F$  already has all derivatives zero at both ends, and thus the transformations do not improve the errors.



Table 7. Fractional errors  $\times 10^5$  for  $F = \exp(-x^{-1} - (1-x)^{-1})$ 

$n$	$T_1$	$T_3$
6	-361.2 $(-1.300)_2$	-1688.7
8	6.344 $(406.0)_2$	107.8
10	20.17 $(0.2017)_2$	- 5.698
12	4.436 $(0.0638)_2$	7.623
16	- 0.7004 $(-0.00179)_2$	0.2687
20	- 0.0379 $(-0.00015)_2$	0.04303
40	0	0
60	0	0
100	0	0

In Table 6, the integral is in fact  $\int_0^1 e^{-r} dr$ , transformed by  $r = x(1-x)^{-1}$  to bring its range to  $[0, 1]$ . Then though all derivatives are zero at  $x = 1$  substantial improvement is still obtained, as would be expected, since the transformations reduce the derivatives at  $x = 0$ .

It is of interest to note that the transformation  $T_\infty$  gives results as accurate as the others from  $n = 16$  upwards, and at  $n = 40$ , is giving the best results, all results being in the region of round-off error.

The small point that, since in the Eq. (4), in the practical applications

$$0.5(G(0) + G(1)) = 0 \quad (5)$$

there are only  $(n-1)$  points of evaluation for  $n$  intervals, can become more important in many dimensions where only  $(n-1)^M$  points are required.

The advantages and simplicity of this procedure as a first step, possible for one dimensional, but very markedly for many dimensions, does not appear to have been widely appreciated. It appears quite possible that combinations of this procedure with diophantine or similar systems of points [Haselgrove (1961), Conroy (1967)], might be one of the most effective methods for difficult many dimensional integrals. The work of Sag and Szekeres could probably be regarded as a special procedure of this type. In the diophantine approach, it is uncertain how much developments to improve accuracy have in the past been expended against boundary errors, more easily removed here, and how much against the interior integration. The freedom to adjust more effectively for just the second source of error might improve such methods considerably. It is considered there are two main sources of error: the boundary corrections, here removed; and the effects of integrating bulges in the interior by finite points. The result for the narrow gaussian curve (Table 5) gives a measure of the errors due to a bulge of width 0.2 and could be scaled to give a rough estimate for any suspected interior bulge effect.

It may be noted that such methods, especially  $T_\infty$ , can be used to integrate functions which have integrable singularities, e.g.

$$x^{-0.5} + x^{-0.8} + (1-x)^{-0.95}.$$

Simple example of such cases are sometimes cited as being appropriate to Gauss point quadrature. But this might necessitate elaborate calculation of points for

each integral. In other simpler cases Gauss point quadrature gives high accuracy but requires rather stringent knowledge that the integrand is very similar to a moderate finite expansion in terms of a particular type of function. This and other aspects practically exclude any general application of Gauss point procedures to more than one or two dimensions.

The above tables provide a means of estimating the errors for given points per dimension in many dimensional integrals after the boundary reduction transformations have been applied. In the simplest application the latter provide a very effective and simple way of reducing the integration to the evaluation of  $(n-1)$  equally spaced points per dimension. It is to be hoped that further efficiency will be obtained by use of these transformations with other improved point systems.

### References

1. Boys, S. F., Handy, N. C.: *Proc. Roy. Soc. (London)* A 311, 309 (1969)
2. Boys, S. F., Rajagopal, P.: *Adv. Quantum Chem.* 2, 1 (1965)
3. Conroy, H.: *J. Chem. Phys.* 47, 5307 (1967)
4. Haselgrove, C. B.: *Math. Comp.* 15, 323 (1961)
5. Krylov, V. I.: *Approximate calculation of integrals*, p. 215. New York: MacMillan 1962
6. Sag, T. W., Szekeres, G.: *Math. Comp.* 18, 245 (1964)
7. Whittaker, E. T., Robinson, G.: *Calculus of observations*, 2nd ed., p. 134. London: Blackie 1937

Dr. N. C. Handy  
Theoretical Chemistry Department  
University Chemical Laboratory  
Lensfield Road  
Cambridge, CB2 1EW  
England

# A Method for Selection and Use of Numerical Integration Points for Molecules with Cylindrical Symmetry\*

N. C. Handy

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1 EW

Received March 30, 1973

Boys and Handy [1] have discussed the solution of the bivariational equations with restricted numerical integration. One of the weaknesses of the method was that in the numerical summations over points, some points arose with  $r_{ij}=0$  and non-zero weights. This makes the method quite impractical for the Schrodinger Hamiltonian (because of the singularity at  $r_{ij}=0$ ), and it cannot be advantageous for the transcorrelated Hamiltonian  $C^{-1}HC$  because there will be some discontinuous higher derivatives at  $r_{ij}=0$ . Here it is shown how the symmetry of cylindrically symmetric molecules can be used to eliminate such points, without losing any of the advantages of the overall method, such as the convergence of the eigensolutions. It is also shown how the primary numerical integration points  $(z_i, r_i)$  may be chosen in any calculation such that each is associated with an equal amount of one-electron density. The choice of the angular coordinates are governed by the removal of the  $r_{ij}=0$  points and maintaining the natural orthogonality between orbitals of different symmetry types. The method has been programmed and found to be practical, although no new molecular calculations have yet been performed. It is to be hoped that these points will give a basis for new transcorrelated calculations on diatomic molecules.

**Key words:** Transcorrelated method Numerical integration

## 1. Removal of $R_{ij}=0$ Contributions

Boys and Handy [1] have discussed the solution of the bivariational equations with restricted numerical integration

$$\langle \Psi, Q | H - W | \sum_i Y_i \Phi_i \rangle = 0 \quad (1)$$

Here  $\Psi$ , and  $\Phi_i$  denote Slater determinants and  $H$  may denote either the Schrodinger Hamiltonian  $H_s$  or the Transcorrelated Hamiltonian  $C^{-1}H_sC$ , and  $W$  is the associated eigenvalue.  $Q$  denotes the restricted numerical integration operator

$$Q = \sum_i^M h_i \delta(r_1, R_i) \sum_j^M h_j \delta(r_2, R_j) \dots \sum_u^M h_u \delta(r_N, R_u) \quad (2)$$

where  $(h_i, R_i)$  are a set of  $M$  weights and points in 3 dimensions, and  $N$  is the number of electrons.

The above authors have underlined the advantages of working with Eq. (1): firstly, the error in  $W$  is proportional to  $\mu$ , the least squares error of the basis set  $\Phi_i$  to the true eigenfunction, whatever the functions  $\Psi$ , and  $Q$ ; secondly,

\*This paper was presented during the session on numerical integration methods for molecules of the 1970 Quantum Theory Conference in Nottingham. It has been revised in the light of the interesting discussion which followed.

a Slater determinant Projective Reduction theorem can be derived for the evaluation of the matrix elements, on a parallel with the standard theorem for the accurate determination of these elements (when  $Q = 1$ ) through 6-dimensional integrals. Thus for a two electron operator  $F$ ,  $\langle \Psi, Q | \sum_{ij}^* F_{ij} | \Phi_s \rangle$  is a linear combination of sums  $S$  of the form

$$S = \sum_i^M h_i \sum_j^M h_j \psi_a^*(R_i) \psi_b^*(R_j) F(R_i, R_j) \varphi_c(R_i) \varphi_d(R_j) \quad (3)$$

where  $\psi$  and  $\varphi$  are one electron orbitals occurring in  $\Psi$ , and  $\Phi_s$ , and satisfying

$$\sum_i h_i \psi_a^*(R_i) \varphi_c(R_i) = \delta_{ac} \quad (4)$$

This procedure was successfully used in calculations with the transcorrelated method on LiH.

In the sum  $S$ , it will be noticed that  $i$  equals  $j$   $M$  times and so  $R_{ij} = 0$  occurs  $M$  times with non zero weights  $h_i^2$ , thus making the method particularly unsuitable for operators which have singularities or discontinuities at  $r_{ij} = 0$ . It is therefore impossible to work with  $H_s$ , because of  $r_{ij}^{-1}$ , and  $C^{-1}H_sC$  may give difficulty because there are certainly discontinuities in higher derivatives at  $r_{ij} = 0$ . This is particularly relevant in the present circumstances when the proportion  $(1/M)$  of these points to the total points is quite significant,  $M$  being small (say 100).

Here it is shown how the symmetry of cylindrically symmetric molecules can be used to remove these undesirable points  $R_i = R_j$ , through a slight redefinition of  $Q$ , the above mentioned advantages still holding true. Essentially the method depends upon the fact that if

$$R_i = (z_i, r_i, \theta_i) \text{ and } R_i^k = \left( z_i, r_i, \theta_i + k \frac{2\pi}{N} \right), \quad k = 1, 2, \dots, N \quad (5)$$

then whenever (4) holds, so also does

$$\sum_i h_i \psi_a^*(R_i^k) \varphi_c(R_i^k) = \delta_{ac} \quad (6)$$

because of the symmetry of the orbitals  $\psi$  and  $\varphi$  around the  $z$  axis. If the operator  $Q$  is therefore replaced by the (symmetric) operator

$$Q = \frac{1}{N!} \sum_k^{N!} P_k \left[ \sum_i^M h_i \delta(r_1, R_i^{k_1}) \sum_j^M h_j \delta(r_2, R_j^{k_2}) \sum_u^M h_u \delta(r_N, R_u^{k_N}) \right] \quad (7)$$

where  $P_k$  runs over all permutations  $k$  of  $1, 2, 3 \dots N$ , then a Slater determinant Projective Reduction theorem can still be derived, because its derivation depends upon the use of the orthogonality conditions Eq. (6). In this case it appears that  $S$  takes the form

$$S = N^{-1} (N-1)^{-1} \sum_{\lambda \neq \mu}^N \sum_i^M h_i \sum_j^M h_j \psi_a^*(R_i^\lambda) \psi_b^*(R_j^\mu) F(R_i^\lambda, R_j^\mu) \varphi_c(R_i^\lambda) \varphi_d(R_j^\mu) \quad (8)$$

where it is immediately noticed that no points ever arise with  $R_{ij} = 0$ .

## 2. Further Substantial Reductions through Symmetry

The cylindrical symmetry can be further used to simplify  $S$ . [For expediency, it is assumed here that the orbitals are only of  $\sigma$  type and that  $F$  is a function of  $z$ ,  $r$  and  $r_{ij}$ ]. It is certainly possible to put all the original points on  $\theta=0$ , that is  $\theta_i=0$ , — Eq. (4) will still hold for  $\sigma$  type orbitals, (for higher symmetry types this is not so, but it is simply corrected by associating with each  $(z_i, r_i)$  a certain number of  $\theta$  points equally spaced around the  $z$  axis). Furthermore, for each term in  $S$ , it is always possible to rotate the  $\theta$  axis such that all the  $R_i$  points lie on  $\theta=0$ , without changing its value, again because of the symmetry. These further arguments give  $S$  in the form

$$S = N^{-1}(N-1)^{-1} \sum_{i \neq \mu}^N \sum_{j \neq \lambda}^M h_i h_j \psi_a^*(R_i^0) \psi_b^*(R_j^{\mu-\lambda}) F(R_i^0, R_j^{\mu-\lambda}) \varphi_c(R_i^0) \varphi_d(R_j^{\mu-\lambda}) \quad (9)$$

where

$$R_i^0 = (z_i, r_i, 0), \quad R_j^{\mu-\lambda} = \left( z_j, r_j, (\mu-\lambda) \frac{2\pi}{N} \right)$$

The operator  $F$  will have different values for different values of  $\left[ (\mu-\lambda) \frac{2\pi}{N} \right] \bmod(\pi)$ , so in effect  $S$  is a sum of  $pM^2$  quantities, where  $p$  is the number of these distinct values. The method is simply extended, as indicated above, to the case when  $\psi$  and  $\varphi$  are of higher symmetry types,  $m_1$  and  $m_2$  say, where the relation (4) takes the form

$$\sum_i h_i \psi_a^{m_1}(R_i) \varphi_c^{m_2}(R_i) = \delta_{ac} \delta_{m_1 m_2} \quad (10)$$

For a calculation on HF involving  $\sigma$ ,  $\pi$ ,  $\delta$  occupied and excited orbitals, which uses all the above procedures, thus maintaining the advantages of the restricted integration bivariational method Eq. (1), yet has no  $R_{ij}=0$  points and also obeys Eq. (10),  $S$  is a sum of  $14M^2$  finite quantities. [Part of the theory used here will be given in detail in a forthcoming publication [2]].

## 3. Selection of the $M$ Basic Points $R_i^0$

It will be assumed for the molecules under consideration that an approximate one electron density function  $\rho(z, r)$  is known, from an SCF calculation or otherwise. The idea here is to divide the 2 dimensional space  $(z, r)$  into regions such that each point is associated with an equal amount of electron density. Such a method can be described as follows:

$$\text{Let } f(z) = \int_0^\infty 2\pi r \rho(z, r) dr \quad \text{and} \quad A = \int_{-\infty}^\infty f(z) dz \quad (11)$$

First divide the  $z$  axis as follows by choosing  $\bar{z}_i$  such that

$$\int_{-\infty}^{\bar{z}_i} f(z) dz = (i/L) A \quad \text{and} \quad z_i = 1/2 (\bar{z}_i + \bar{z}_{i-1}) \quad (12)$$

where  $i = 1, 2, \dots, L$ , and  $L \simeq M^{\frac{1}{2}}$ ,  $M$  being the total number of points required. For each  $z_i$  point now determined, the values  $r_i^j$  associated with it are similarly selected. Let

$$\int_0^\infty 2\pi r \rho(z_i, r) dr = (j/L) f(z_i) \quad \text{and} \quad r_i^j = \frac{1}{2} (\bar{r}_i^j + \bar{r}_i^{j-1}) \quad (13)$$

for  $j = 1, 2, \dots, L$ .

As the problem here is only 2 dimensional, all the integrals in the above expressions can be accurately evaluated numerically, using a sufficiently large number of points. Thus in practice  $\pm \infty$ ,  $\bar{z}_0$  and  $\bar{r}_i^0$  are each replaced by some suitably large value. The  $L^2 (\simeq M)$  points  $(z_b, r_i)$  selected by this procedure will each have equal weight, the 1-electron density space having been divided into  $L^2$  equal areas. In a typical transcorrelated calculation on HF, it is anticipated that  $M \simeq 144$ ,  $L \simeq 12$ .

#### 4. Conclusion

The above procedures have been programmed for diatomic molecules, and have been found entirely practical. The increase in the number of 6 dimensional points from  $M^2$  to  $pM^2$  does not increase the time factor by  $p$ , because of many simplifying details — and in any case, this factor is surely warranted because of the removal of the  $R_{ij}=0$  points.

#### References

1. Boys, S. F., Handy, N. C.: Proc. Roy. Soc. (London) A 311, 309 (1969)
2. Boys, S. F., Handy, N. C.: To be published

Dr. N. C. Handy  
Theoretical Chemistry Department  
University Chemical Laboratory  
Lensfield Road  
Cambridge CB2 1EW  
England

## SCF LCGO MO Studies on the Fluoronium Ion $\text{FH}_2^+$ and Its Hydrogen Bonding Interaction with Hydrogen Fluoride FH

G. H. F. Diercksen, W. von Niessen, and W. P. Kraemer

Max-Planck-Institut für Physik und Astrophysik, 8000 München 40, Germany

Received February 26, 1973

The stability and geometrical structure of the fluoronium ion is investigated using the one-determinant SCF LCAO MO method. The equilibrium geometry is characterized by a bond length of  $d(\text{FH}) = 0.95 \text{ \AA}$  and a bond angle of  $114.75^\circ$ . The proton binding energy is determined to be 120.1 kcal/mole. The molecules  $\text{FH}_3^+$  and  $\text{FH}_3$  are found to be unstable. A binding energy of 30.7 kcal/mole is obtained for the hydrogen bond formation between the systems  $\text{FH}_2^+$  and FH. In the minimum energy structure the central proton is situated midway between the two F atoms in a symmetrical single minimum potential. The general behavior of the potential curves of the di-solvated proton involving  $\text{NH}_3$ ,  $\text{OH}_2$ , and FH as solvent molecules is discussed. In all these cases double minimum potentials are found, if the equilibrium separation between the heavy atoms is larger than approximately  $2.4 \text{ \AA}$ , and single minimum potential for separations smaller than this value.

**Key words:** SCF LCGO MO calculation – Proton solvation    Hydrogen bonding –  $\text{FH}_2^+$   
 $\text{FH}_2^+ \cdot \text{FH}$

The present investigation deals with two topics, the solvation of a proton by a single hydrogen fluoride molecule and the hydrogen bonding interaction between the resulting fluoronium ion and another hydrogen fluoride molecule. While the ammonium and oxonium ions  $\text{NH}_4^+$  and  $\text{OH}_3^+$  are well established experimentally, this has not been possible until recently [1] for the fluoronium ion  $\text{FH}_2^+$ , although its existence has been postulated to explain in particular the electrical conductivity of acid solutions of several fluoride compounds in liquid hydrogen fluoride. Further, from the high binding energies of the proton attached to an ammonia or a water molecule, which are experimentally determined to be  $B(\text{NH}_4^+) = 207 \text{ kcal/mole}$  [2] and  $B(\text{OH}_3^+) = 168 \text{ kcal/mole}$  [3], it can be concluded that the fluoronium ion forms a relatively stable species too. The corresponding proton binding energy is expected to be somewhat smaller than  $B(\text{OH}_3^+)$  due to the higher electronegativity of the fluorine atom as compared to the oxygen atom.

In a recent publication Couzy and co-workers reported the first direct observation of the fluoronium ion [1]. By an analysis of the infrared spectra of some fluoronium salts in the solid and liquid state the geometrical structure of the  $\text{FH}_2^+$  ion was shown to be angular. But so far no accurate experimental determination of the geometrical parameters of the  $\text{FH}_2^+$  ion has been performed.

So the fluoronium ion seems to provide an ideal case for a theoretical prediction of the geometrical structure because of the small size of the electronic system. Apart from a quite simplified model calculation (FSGO model) by Frost [4] the only theoretical studies on the  $\text{FH}_2^+$  system known to the authors are LCAO MO SCF calculations by Csizmadia and co-workers with various size gaussian basis sets [5] and semi-empirical calculations of Schuster *et al.* using the CNDO SCF method followed by a configuration interaction expansion (CNDO CI) [6]. Only in these latter semiempirical studies the optimum geometrical structure of  $\text{FH}_2^+$  was searched, while in Csizmadia's SCF calculations a fixed geometry of the nuclear centres has been assumed ( $d(\text{FH}) = 0.917 \text{ \AA}$ ,  $\angle(\text{HFH}) = 105^\circ$  [7]). In the present work it was tried to get accurate theoretical informations about the geometrical structures of the ionic systems  $\text{FH}_2^+$  and  $\text{FH}_2^+ \cdot \text{FH}$  within the one-determinant LCAO MO SCF framework using a rather extended gaussian basis set. For the system  $\text{FH}_2^+ \cdot \text{FH}$  in particular the potentials of the central proton, involved into the hydrogen bond formation between the two FH fragments has been computed for different F-F distances. These potential curves are discussed and compared to those obtained for the systems  $\text{NH}_4^+ \cdot \text{NH}_3$ ,  $\text{OH}_3^+ \cdot \text{OH}_2$ , and  $\text{OH}_2 \cdot \text{OH}^-$ .

The SCF wavefunctions and energy expectation values have been calculated using Roothaan's SCF LCAO MO expansion method [8]. The calculations were carried out on an IBM 360/91 with the program system MUNICH [9], which is based on the use of general Gaussian functions  $\eta = x^l y^m z^n \exp(-\alpha r^2)$  as basis functions for the expansion of the molecular orbitals. A (11s7p1d/6s1p) basis set was employed for the fluorine and hydrogen centres, respectively, contracted to a [5s4p1d/1s1p] set to reduce the number of linear parameters in the SCF procedure (for a definition of the basis set notation used here, see [10]). The exponential parameters and contraction coefficients are taken from the literature [11]. The polarization functions (d-type functions on the F atom and p-type functions on the H atom) were optimized by SCF calculations on the FH molecule and their exponential parameters are obtained to be  $\alpha_d(\text{F}) = 1.23$  and  $\alpha_p(\text{H}) = 0.75$  [12]. With this basis set an SCF energy of  $E^{\text{SCF}} = -100.05638$  a.u. was calculated for FH at its experimental bond length of 1.7329 a.u. This compares with the best SCF energy value reported in the literature of  $E^{\text{SCF}} = -100.07040$  a.u. [13], which is believed to be very close to the Hartree-Fock limit. The energy value obtained with the basis set given above thus differs from this limit by about 0.015 a.u. (ca. 9.5 kcal/mole). The energy effects to be discussed here are much higher than this difference. The basis set used can consequently be regarded to yield reliable results for the geometrical structures and binding energies of the systems to be investigated within the accuracy of the Hartree-Fock method.

As mentioned previously the experimental work on the fluoronium ion  $\text{FH}_2^+$  did not permit an accurate determination of its structure. The ion was supposed to have a bond length of 1.02 Å and a bond angle somewhat larger than the angle in the water molecule (a range of 105–120° was regarded as probable) [1]. Optimization of these geometrical parameters in the present SCF study gave a significantly different bond length of  $d(\text{FH}) = 1.788$



Table 1. Total SCF energies for  $\text{FH}_2^+$  (variation of bond length  $d(\text{FH})$  and  $\text{HFH}$  angle; all values in atomic units)

	$d(\text{FH})$	$\angle \text{HFH}$	$E^{\text{SCF}}$
1	1.93755	$105^\circ$	-100.23632
2	1.92755	$105^\circ$	-100.23746
3	1.91755	$105^\circ$	-100.23854
4	1.90755	$105^\circ$	-100.23956
5	1.88755	$105^\circ$	-100.24140
6	1.86755	$105^\circ$	-100.24296
7	1.84755	$105^\circ$	-100.24421
8	1.82755	$105^\circ$	-100.24515
9	1.80755	$105^\circ$	-100.24573
10	1.79755	$105^\circ$	-100.24588
11	1.78755	$105^\circ$	-100.24593
12	1.77755	$105^\circ$	-100.24588
13	1.76755	$105^\circ$	-100.24572
14	1.74755	$105^\circ$	-100.24507
15	1.78755	$103^\circ$	-100.24504
16	1.78755	$107^\circ$	-100.24664
17	1.78755	$110^\circ$	-100.24737
18	1.78755	$114^\circ$	-100.24777
19	1.78755	$114.75^\circ$	-100.24778
20	1.78755	$115^\circ$	-100.24778
21	1.78755	$116^\circ$	-100.24775
22	1.78755	$120^\circ$	-100.24727
23	1.78755	$140^\circ$	-100.23889
24	1.78755	$180^\circ$	-100.22309
25	1.77755	$114^\circ$	-100.24775
26	1.77755	$115^\circ$	-100.24776
27	1.77755	$116^\circ$	-100.24773
28	1.79755	$114^\circ$	-100.24769
29	1.79755	$115^\circ$	-100.24770
30	1.79755	$116^\circ$	-100.24766

a.u. =  $0.95 \text{ \AA}$  and a bond angle of  $\angle(\text{HFH}) = 114.75^\circ$  yielding a total SCF energy of  $E^{\text{SCF}} = -100.24778$  a.u. The results of the corresponding calculations are listed in Table 1. Schuster *et al.* calculated a FH bond length of  $d(\text{FH}) = 1.956$  a.u. =  $1.04 \text{ \AA}$  and a bond angle of  $\angle(\text{HFH}) = 120^\circ$  with the CNDO SCF method. Including CI they obtained  $d(\text{FH}) = 1.977$  a.u. =  $1.05 \text{ \AA}$  and  $\angle(\text{HFH}) = 114.5^\circ$ . This latter result for the bond angle is in good agreement with the ab initio SCF results of the present study, while both CNDO values for the bond length are not quite satisfactory. The proton binding energy corresponding to the reaction  $\text{FH}_2^+ \rightleftharpoons \text{FH} + \text{H}^+$  is determined from the SCF results to be  $B(\text{FH}_2^+) = 120.1$  kcal/mole. The binding energy values obtained from both CNDO calculations [CNDO SCF:  $B(\text{FH}_2^+) = 198$  kcal/mole; CNDO CI:  $B(\text{FH}_2^+) = 166$  kcal/mole] are significantly too high. The obtained proton binding energy in the fluoronium ion of  $B(\text{FH}_2^+) = 120.1$  kcal/mole compares well with the corresponding energy values for the ammonium ion [ $B(\text{NH}_4^+) = 207$  kcal/mole] and the oxonium ion [ $B(\text{OH}_3^+)$ ].

Table 2. Total SCF energies for  $\text{FH}_3^{2+}$  (variation of bond length  $d(\text{FH})$  and out of plane angle  $\phi$ ; all values in atomic units)

	$d(\text{FH})$	$\phi$	$E^{\text{SCF}}$
1	1.78	$0^\circ$	-100.05157
2	1.80	$0^\circ$	-100.05594
3	1.90	$0^\circ$	-100.06957
4	1.96	$0^\circ$	-100.07244
5	1.98	$0^\circ$	-100.07269
6	2.00	$0^\circ$	-100.07263
7	2.02	$0^\circ$	-100.07228
8	2.10	$0^\circ$	-100.06841
9	1.98	$5^\circ$	-100.07223
10	1.98	$10^\circ$	-100.07064
11	1.98	$20^\circ$	-100.06126

= 168 kcal/mole], the binding energy differences being about 40 kcal/mole between  $\text{NH}_4^+$  and  $\text{OH}_3^+$  and between  $\text{OH}_3^+$  and  $\text{FH}_2^+$ .

The high stability of the ion  $\text{FH}_2^+$  and the fact that only two hydrogen centres are bound to the fluorine atom gives rise to the question whether the fluoronium ion is able to bind another proton to form  $\text{FH}_3^{2+}$ . The results of SCF calculations connected with this question are summarized in Table 2. From these data it is seen that the  $\text{FH}_3^{2+}$  system has a planar structure with a FH bond length of  $d(\text{FH}) = 1.98$  a.u. = 1.05 Å, which is increased by ca 11% compared to the bond length in  $\text{FH}_2^+$ . The minimum SCF energy is obtained to be  $E^{\text{SCF}} = -100.07269$  a.u. This means that the di-protonated system  $\text{FH}_3^{2+}$  is stable against a dissociation into  $\text{FH} + 2\text{H}^+$  with a binding energy of  $B = 10.2$  kcal/mole (the average binding energy per bonded proton is thus 5.1 kcal/mole). On the other hand, the  $\text{FH}_3^{2+}$  system is seen to be unstable with respect to a dissociation into  $\text{FH}_2^+ + \text{H}^+$  by 109.9 kcal/mole. This result may easily be understood from purely electrostatical considerations. The repulsional forces between the two protons are too strong to be compensated by the sum of the binding energies of each proton to the fluorine centre in the FH molecule.

Similar stability studies are performed for the neutral  $\text{FH}_3$  molecule. The SCF results are collected in Table 3. As in  $\text{FH}_3^{2+}$  the neutral system is obtained to be planar with a FH bond length of  $d(\text{FH}) = 2.25$  a.u. = 1.19 Å. From the minimum SCF energy value of  $E^{\text{SCF}} = -100.85866$  a.u. it follows that the  $\text{FH}_3$  molecule is unstable with respect to a dissociation into  $\text{FH} + \text{H}_2$  by 206.8 kcal/mole. For this purpose the SCF energy of the  $\text{H}_2$  molecule was calculated to be  $E^{\text{SCF}} = -1.131644$  a.u. at a bond length of  $d(\text{HH}) = 1.4$  a.u. using the same basis set as for the H centres in the present study.

In the following paragraphs the interaction between the fluoronium ion and another hydrogen fluoride molecule is studied in some more detail. The geometrical configuration where the two FH fragments are linked together via a linear hydrogen bond can be considered to be the most stable structure for the ionic system:  $(\text{HFHFH})^+$ . Optimization of the geometrical parameters

Table 3. Total SCF energies for  $\text{FH}_3$  (variation of bond length  $d(\text{FH})$  and out of plane angle  $\phi$ ; all values in atomic units)

	$d(\text{FH})$	$\phi$	$E^{\text{SCF}}$
1	1.78	$0^\circ$	-100.77664
2	1.80	$0^\circ$	-100.78488
3	1.90	$0^\circ$	-100.81779
4	2.10	$0^\circ$	-100.85216
5	2.22	$0^\circ$	-100.85839
6	2.24	$0^\circ$	-100.85862
7	2.25	$0^\circ$	-100.85866
8	2.26	$0^\circ$	-100.85864
9	2.30	$0^\circ$	-100.85813
10	2.25	$5^\circ$	-100.85799
11	2.25	$10^\circ$	-100.85614
12	2.25	$20^\circ$	-100.85061

in this configuration gives a minimum SCF energy of  $E^{\text{SCF}} = -200.35304$  a.u. The geometrical parameters are determined to be  $d(\text{FH}^{\text{centr.}}) = 2.14$  a.u.  $= 1.13 \text{ \AA}$  ( $\text{H}^{\text{centr.}}$  is the central hydrogen atom involved in hydrogen bond formation),  $d(\text{FH}^{\text{end}}) = 1.75$  a.u.  $= 0.93 \text{ \AA}$  ( $\text{H}^{\text{end}}$  are the two hydrogen atoms at the ends of the ion) and  $\angle(\text{H}^{\text{end}}\text{F}\text{H}^{\text{centr.}}) = 120.96^\circ$ . The corresponding SCF results are listed in Table 4. A brief survey of the course of independent optimizations of the various geometrical parameters is presented in Table 5. The italic parameter values have been optimized to the given value with the remaining parameters kept fixed to the values listed in the same line. The minimum SCF energy obtained for the actual optimization is given in the last column.

Table 4. Total SCF energies for  $(\text{HFHFH})^+$  (variation of geometrical parameters  $d(\text{H}^{\text{centr.}}\text{F})$  ( $\text{H}^{\text{centr.}}$  is the H atom in the hydrogen bond),  $d(\text{H}^{\text{end}}\text{F})$  ( $\text{H}^{\text{end}}$  are the outer H atoms), and  $\text{HFH}$  bond angle; all values in atomic units)

	$d(\text{H}^{\text{centr.}}\text{F})$	$d(\text{H}^{\text{end}}\text{F})$	$\angle(\text{HFH})$	$E^{\text{SCF}}$
1	1.78755	1.78755	$114.75^\circ$	-200.29102
2	1.82755	1.78755	$114.75^\circ$	-200.30660
3	2.00755	1.78755	$114.75^\circ$	-200.34499
4	2.12755	1.78755	$114.75^\circ$	-200.35131
5	2.14755	1.78755	$114.75^\circ$	-200.35137
6	2.20755	1.78755	$114.75^\circ$	-200.35025
7	2.14755	1.90	$114.75^\circ$	-200.34107
8	2.14755	1.85	$114.75^\circ$	-200.34690
9	2.14755	1.75	$114.75^\circ$	-200.35224
10	2.14755	1.70	$114.75^\circ$	-200.35083
11	2.14755	1.75	$119.75^\circ$	-200.35299
12	2.14755	1.75	$124.75^\circ$	-200.35273
13	2.14755	1.75	$120.96^\circ$	-200.35301
14	2.12755	1.75	$120.96^\circ$	-200.35299
15	2.16755	1.75	$120.96^\circ$	-200.35281
16	2.13942	1.749	$120.96^\circ$	-200.35304

Table 5. Course of optimization of geometrical parameters in (HFHFH)<sup>+</sup> (all values in atomic units)

$d(\text{H}^{\text{center}}\text{F})$	$d(\text{H}^{\text{end}}\text{F})$	$\alpha(\text{HFH})$	$E^{\text{SCF}}$
2.14755	1.78755	114.75°	- 200.35137
2.14755	1.749	114.75°	- 200.35224
2.14755	1.75	120.96°	- 200.35301
2.13942	1.749	120.96°	- 200.35304

Table 6. Theoretical and experimental binding energies for the systems  $\text{H}^+ \cdot \text{XH}_n$  and  $\text{H}^+ \cdot 2\text{XH}_n$  ( $\text{XH}_n = \text{NH}_3, \text{OH}_2, \text{FH}$ ). (All values in atomic units)

Molecule	$B_1(\text{theor})^a$	$B_1(\text{exp})^a$	$B_{2(\text{av})}(\text{theor})^b$	$B_2(\text{theor})^c$	$B_2(\text{exp})^c$
$\text{H}^+ \cdot \text{NH}_3$	215.84 [11]	207.0 [2]	-	-	-
$\text{H}^+ \cdot \text{OH}_2$	173.05 [12]	168.2 ± 3.4 [3]	-	-	-
$\text{H}^+ \cdot \text{FH}$	120.14	-	-	-	-
$\text{H}^+ \cdot 2\text{NH}_3$	-	-	125.0 [13]	32.0 [13]	-
$\text{H}^+ \cdot 2\text{OH}_2$	-	-	102.7 [12]	32.36 [12]	36.0 [15]
$\text{H}^+ \cdot 2\text{FH}$	-	-	75.35	30.68	-

<sup>a</sup>  $B_1 = E(\text{H}^+ \cdot \text{XH}_n) - E(\text{XH}_n)$ .<sup>b</sup>  $B_{2(\text{av})} = (E(\text{H}^+ \cdot 2\text{XH}_n) - E(\text{XH}_n))/2$ .<sup>c</sup>  $B_2 = E(\text{H}^+ \cdot 2\text{XH}_n) - E(\text{H}^+ \cdot \text{XH}_n) - E(\text{XH}_n)$ .

The bond energy of the molecule (HFHFH)<sup>+</sup> with respect to the total dissociation into 2 FH molecules and a proton follows to be  $B_{\text{total}} = 150.8$  kcal/mole. Subtracting from this  $B_{\text{total}}$ -value the first solvation energy of the proton  $B_1 = B(\text{FH}_2^+) = 120.1$  kcal/mole a hypothetical second solvation energy of  $B_2 = 30.7$  kcal/mole is obtained which corresponds to the reaction:  $\text{H}^+ \cdot 2\text{FH} \rightleftharpoons \text{H}^+ \cdot \text{FH} + \text{FH}$ . On the other hand an average di-solvation energy of  $B_{2(\text{av})} = 75.4$  kcal/mole can be defined from the  $B_{\text{total}}$ -value, corresponding to the di-solvation reaction  $\text{H}^+ \cdot 2\text{FH} \rightleftharpoons \text{H}^+ + 2\text{FH}$ . These energy results may be compared to the first and second solvation energies and to the average di-solvation energies of the proton with respect to ammonia and water as solvent systems. The corresponding values are collected in Table 6 together with experimental data if available. From this comparison it follows that the first solvation energies  $B_1$  decrease with increasing nuclear charge of the central atom in the solvent molecule; i.e. with increasing acid character of the solvent molecule. The same behavior is observed in the series of the average di-solvation energies. It is somewhat surprising that the second solvation energies remain approximately constant within the accuracy of the present SCF calculations. The enormous difference between the first and second solvation energies makes it highly improbable for a third solvent molecule to be directly bonded to the central proton. From experimental data on crystal structures it is known that non-bonded oxygen atoms cannot come closer to each other than about 3.5 Å [19]. Similar results are expected to hold for nitrogen and fluorine atoms. Keeping non-bonded X-atoms (X = N, O, F) at a distance of 3.5 Å apart from each other the distance to the central proton would be  $d(\text{XH}) = 2.02$  Å for an

equilateral triangular arrangement of the X-centres and  $d(\text{XH}) = 2.13 \text{ \AA}$  for a tetrahedral configuration of the X-centres. Thus the triangular and tetrahedral arrangements of the solvent molecules lead to a remarkable increase in the XH-bond length compared to the  $d(\text{XH})$  values obtained for the monosolvate systems  $\text{H}^+ \cdot \text{XH}_n$ . From the resulting loss in XH-bond stability it can be concluded that chain structures are more likely to be formed than solvate clusters with a central proton. In these chains the bond energy per additional hydrogen bond to be formed is expected to approach rapidly the energy value calculated for the H-bond energy between the neutral solvent molecules [i.e. 2.7 kcal/mole for  $(\text{NH}_3)_2$ , 4.9 kcal/mole for  $(\text{OH}_2)_2$ , and 4.5 kcal/mole for  $(\text{FH})_2$ ].

Finally the potentials of the central hydrogen atom have been calculated for the equilibrium configuration of the  $(\text{HFHFH})^+$  ion with a FF-distance of  $d(\text{FF}) = 4.279 \text{ a.u.}$  as well as for a slightly enlarged FF-distance of  $d(\text{FF}) = 4.599 \text{ a.u.}$  The results are listed in Table 7 and displayed in Fig. 1. For the equilibrium configuration the potential curve has a symmetrical single minimum; i.e. the optimum position of the central hydrogen atom is midway between the two F atoms. At the larger FF-distance of  $d(\text{FF}) = 4.599 \text{ a.u.}$  a symmetrical double minimum potential results. The height of the barrier is only 0.7 kcal/mole. At the two minima of the potential curve the central H atom is off from the midway position by about 0.3 a.u. Similar calculations have been performed for the systems  $(\text{H}_2\text{OHOH}_2)^+$  [15],  $(\text{HOHOH})^-$  [20], and  $(\text{H}_3\text{NHNH}_3)^+$  [16, 17]. The potential curves for the oxonium hydrate ion and the hydroxyl hydrate ion are displayed for comparison in Fig. 2 and 3. The equilibrium configurations of the systems involve a double minimum potential curve for  $(\text{HOHOH})^-$  and  $(\text{H}_3\text{NHNH}_3)^+$ , an indication of a double minimum for  $(\text{H}_2\text{OHOH}_2)^+$  where the barrier is beyond the numerical interpretability, and a single minimum curve for  $(\text{HFHFH})^+$ . Comparing the potential curves for the various distances  $d(\text{XX})$  between the two heavy atoms of the systems  $(\text{H}_n\text{XHXH}_n)^\pm$  one can draw

Table 7. Total SCF energies for the proton transfer in  $(\text{HFHFH})^+$  at two different FF separations. (All values in atomic units)

1	$d(\text{FF})$	$\Delta(\text{H}^{\text{centr}})^a$	$E^{\text{SCF}}$
1	4.27884	0.0	-200.35304
2	4.27884	0.1	-200.35271
3	4.27884	0.2	-200.35152
4	4.27884	0.4	-200.34299
5	4.59884	0.0	-200.34695
6	4.59884	0.1	-200.34717
7	4.59884	0.2	-200.34769
8	4.59884	0.3	-200.34808
9	4.59884	0.4	-200.34756
10	4.59884	0.5	-200.34495
11	4.59884	0.6	-200.33855

<sup>a</sup>  $\Delta(\text{H}^{\text{centr}})$  gives the displacement of the proton from the midway position between the two F atoms.

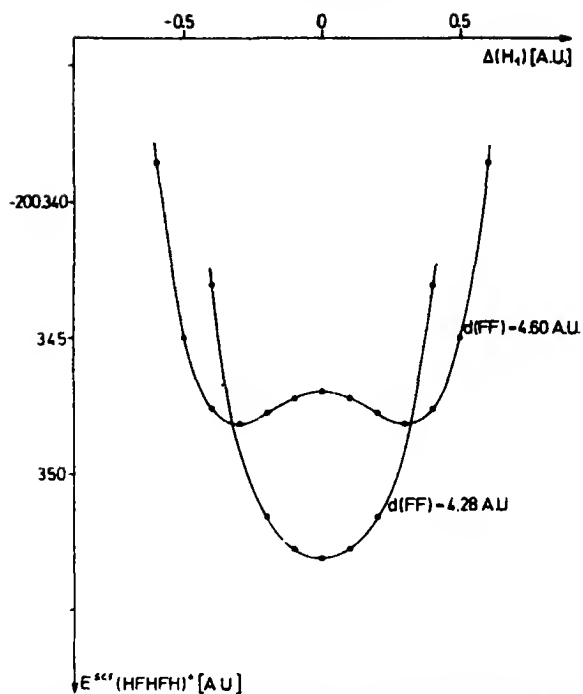


Fig. 1. Potential energy curves for the proton transfer in  $(\text{HFHFH})^+$  at two different F-F separations [ $d(\text{FF}) = 4.28$  a.u. and  $d(\text{FF}) = 4.60$  a.u.]. (The abscissa gives the displacement of the proton from the midway position between the two F atoms)

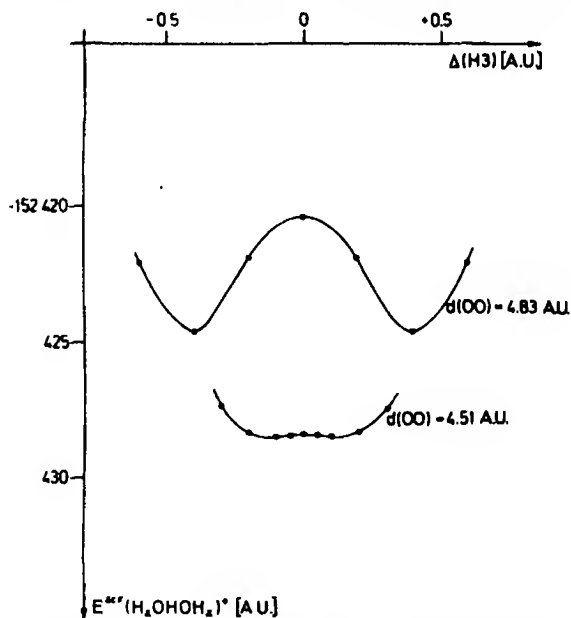


Fig. 2. Potential energy curves for the proton transfer in  $(\text{H}_2\text{OHOH}_2)^+$  at two different O-O separations [ $d(\text{OO}) = 4.51$  a.u. and  $d(\text{OO}) = 4.83$  a.u.]. (The abscissa gives the displacement of the proton from the midway position between the two O atoms)

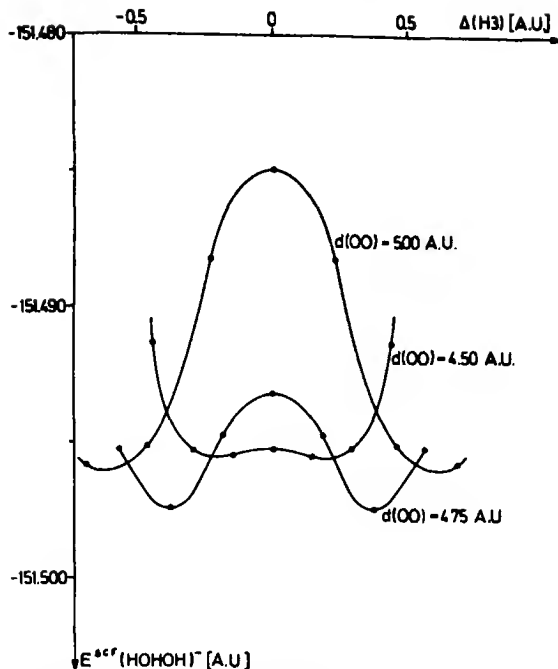


Fig. 3. Potential energy curves for the proton transfer in  $(\text{HOHOH})^-$  at three different O—O' separations [ $d(\text{OO}) = 4.50$  a.u.,  $d(\text{OO}) = 4.75$  a.u., and  $d(\text{OO}) = 5.00$  a.u.]. (The abscissa gives the displacement of the proton from the midway position between the two O atoms)

the following conclusions. If  $d(\text{XX})$  is smaller than approximately  $4.5 \text{ a.u.} = 2.4 \text{ \AA}$  a single minimum potential curve results for the central hydrogen atom, whereas for  $d(\text{XX})$ -values larger than  $4.5 \text{ a.u.}$  a double minimum potential is found. For  $d(\text{XX})$ -values in the neighbourhood of  $4.5 \text{ a.u.}$  the barrier begins to develop. In the series of ionic systems studied here this  $d(\text{XX})$ -value seems to be relatively insensitive to the nature of the heavy atoms in the solvent molecules and to whether the molecular systems is a positive or a negative ion. Although it is not possible to make general conclusions from the limited number of cases considered, this result seems to be quite remarkable for hydrogen bond formation in ionic molecules.

*Acknowledgements.* It is a pleasure to thank our technical staff for careful preparation of the input for the programs and for its skillful assistance in running the computer.

### References

1. Couzi, M., Corunt, J.C., Huong, P.V.: *J. Chem. Phys.* **56**, 426 (1972)
2. Haney, M.A., Franklin, J.L.: *J. Chem. Phys.* **50**, 2028 (1969)
3. Chong, S.L., Myers, R.A., Franklin, J.L.: *J. Chem. Phys.* **56**, 2427 (1972)
4. Frost, A.A.: *J. Phys. Chem.* **72**, 1289 (1968)
5. Hopkinson, A.C., Holbrook, N.K., Yates, K., Csizmadia, I.G.: *J. Chem. Phys.* **49**, 3596 (1968)
6. Jakubetz, W., Lischka, H., Rosmus, P., Schuster, P.: *Chem. Phys. Letters* **11**, 38 (1971)

7. Kuipers, G. A., Smith, D. F., Nielsen, A. H.: *J. Chem. Phys.* **25**, 275 (1956)
8. Roothaan, C. C. J.: *Rev. Mod. Phys.* **23**, 69 (1951)
9. Dierksen, G. H. F., Kraemer, W. P.: MUNICH, Molecular Program System, Reference Manual, Special Technical Report, Max-Planck-Institut für Physik und Astrophysik (to be published)
10. Moskowitz, J. W., Harrison, M. C.: *J. Chem. Phys.* **43**, 3550 (1965)
11. Salez, C., Veillard, A.: *Theoret. chim. Acta (Berl.)* **11**, 441 (1968)
12. Dierksen, G. H. F., Kraemer, W. P.: *Chem. Phys. Letters* **6**, 419 (1970)
13. McLean, A. D., Yoshimine, M.: *J. Chem. Phys.* **47**, 3256 (1967). —  
Cade, P. E., Huo, W. M.: *J. Chem. Phys.* **47**, 614 (1967)
14. Dierksen, G. H. F.: unpublished results
15. Kraemer, W. P., Dierksen, G. H. F.: *Chem. Phys. Letters* **5**, 463 (1970), and unpublished results
16. Merlet, P., Peyerimhoff, S. D., Buenker, R. J.: preprint 1972
17. Delpuech, J. J., Serratrice, G., Strich, A., Veillard, A.: *Chem. Com.* to be published
18. Kebabian, P., Searles, S. K., Zolla, A., Scanborough, J., Arshadi, M.: *J. Am. Chem. Soc.* **89**, 6393 (1967)
19. See e.g.: Glasstone, S.: *Textbook of physical chemistry*, p. 384. London: McMillan, 1956
20. Kraemer, W. P., Dierksen, G. H. F.: *Theoret. chim. Acta (Berl.)* **23**, 398 (1972)

G. H. F. Dierksen  
Max-Planck-Institut  
für Physik und Astrophysik  
D-8000 München 40  
Föhringer Ring 6  
Federal Republic of Germany



# CNDO/2 (Complete Neglect of Differential Overlap)- Method for Third-Row Molecules

H. L. Hase and A. Schweig

Fachbereich Physikalische Chemie der Universität, D-3550 Marburg (Lahn), Germany

Received February 23, 1973

An extension of the CNDO/2 method to compounds containing third-row elements (Germanium, Arsenic, Selenium and Bromine) is presented. Bond lengths, bond angles, dipole moments, and ionization potentials are considered.

**Key words:** CNDO/2-method for third-row atoms – All-valence electron calculations – Dipole moment formula for third-row molecules

## 1. Introduction

The currently available CNDO/2 computer program [1] based on the CNDO/2 (Complete Neglect of Differential Overlap)-method [2–8] permits valence electron calculations of molecules containing hydrogen, and first- and second-row atoms. We have now extended the method and corresponding program to third-row elements. A previous extension [9–14] did not follow the original CNDO/2 parametrization scheme.

## 2. Method

The elements of the CNDO/2 energy matrix are [5]:

$$F_{\mu\mu} = -\frac{1}{2}(I_{\mu} + A_{\mu}) + [(P_{AA} - Z_A) - \frac{1}{2}(P_{\mu\mu} - 1)]\gamma_{AA} + \sum_{B \neq A} (P_{BB} - Z_B)\gamma_{AB}, \quad (1)$$

$$F_{\mu\nu} = \beta_{AB}^0 S_{\mu\nu} - \frac{1}{2}P_{\mu\nu}\gamma_{AB}, \quad (2)$$

where  $\mu$  and  $\nu$  refer to two valence atomic orbitals (AOs)  $\phi_{\mu}$  and  $\phi_{\nu}$ . All valence AOs (*s*-, *p*- and *d*-AOs) are considered.  $I_{\mu}$  and  $A_{\mu}$  are the ionization potential and electron affinity of  $\phi_{\mu}$ .  $P_{\mu\mu}$  and  $P_{\mu\nu}$  are elements of the charge density matrix (number of electrons which populate  $\phi_{\mu}$  and the overlap region  $\phi_{\mu}\phi_{\nu}$ , respectively) and  $S_{\mu\nu}$  elements of the overlap matrix.  $P_{AA}$  ( $P_{BB}$ ) refers to the total number of valence electrons on atom *A* (*B*) [found by summing all  $P_{\mu\mu}$  on atom *A* (*B*)] and  $Z_A$  ( $Z_B$ ) to the total number of valence electrons contributed by atom *A* (*B*).  $\gamma_{AB}$  ( $\gamma_{AA}$ ) denotes the repulsion integrals between two electrons both in *s* AOs, one on atom *A* and the other on atom *B* (both in the same *s* AO on atom *A*).  $\beta_{AB}^0$  is a bonding parameter.

The extension of the CNDO/2 method to third-row molecules involves five steps.

### 2.1. Choice of Valence AOs for Third-Row Elements

As for the preceding row elements, we chose as basis AOs Slater-type orbitals (STOs) [15]. Their radial part is [5, 16]

$$R_{n,l} = (2\xi)^{n+\frac{1}{2}} [(2n)!]^{-\frac{1}{2}} r^{n-1} e^{-\xi r}, \quad (3)$$

where  $n$  is the principal quantum number and  $\xi = (Z - s)/n^*$  the orbital exponent ( $Z$  is the atomic number,  $s$  a screening constant determined by Slater's rules [15] and  $n^*$  an effective principal quantum number). The same orbital exponent is given to  $4s$ ,  $4p$  and  $4d$  AOs (*spd* basis set). Its numerical values are: Ge = 1.527, As = 1.702, Se = 1.878, and Br = 2.054.

2.2. Evaluation of Overlap Integrals  $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$ , where  $\phi_\mu$  or  $\phi_\nu$  Refers to  $4s$ ,  $4p$  or  $4d$  AOs and Repulsion Integrals  $\gamma_{AB} = \langle s_A^2(1) | e^2/r_{12} | s_B^2(2) \rangle$ , where  $s_A$  or  $s_B$  Denotes a  $4s$  AO

In the original CNDO/2 program [1] provision was made for evaluation of overlap and repulsion integrals involving AOs of third-row elements. Thus after the determination of several new expansion coefficients [17] the calculation of the integrals can be achieved without difficulty.

2.3. Choice of Orbital Electronegativities  $\frac{1}{2}(I_\mu + A_\mu)$  for  $4s$ ,  $4p$ , and  $4d$  AOs

The orbital electronegativities [18] of  $4s$  and  $4p$  AOs are obtained by a previously described method [4, 19]. They are (in eV) for  $4s$  AOs: Ge = 11.435, As = 13.335, Se = 16.315 and Br = 19.630 and for  $4p$  AOs: Ge = 4.080, As = 5.345, Se = 7.100 and Br = 8.400. Values for  $4d$  orbitals are assumed to be zero [7].

### 2.4. Empirical Determination of Bonding Parameters $\beta_{AB}^0$

The bonding parameters  $\beta_{AB}^0$  are written in terms of the atomic parameters  $\beta_A^0$  and  $\beta_B^0$ :

$$\beta_{AB}^0 = \frac{1}{2} K (\beta_A^0 + \beta_B^0) \quad (4)$$

Here  $\beta_A^0$  or  $\beta_B^0$  refer to third-row atoms.  $K$  is given the value 0.75 [5]. The best atomic parameters obtained from an optimization of bond lengths, bond angles, dipole moments, and ionization potentials are (in eV): Ge = -10, As = -13, Se = -16, and Br = -22.

### 2.5. Derivation of the Formula for Calculation of Dipole Moments of Molecules Containing Third-Row Elements

Three terms contribute to the molecular dipole moment [3, 5-7]:  $\mu_Q$  due to the net atomic charges,  $\mu_{sp}$  due to the *sp* atomic polarization (mixing of *s* and *p* AOs), and  $\mu_{pd}$  due to *pd* mixing. The Cartesian components of the three contributions

$[\mu_Q(x)$  and  $\mu_Q(y)$  are of the same form as  $\mu_Q(z)$ ; the same is true for  $\mu_{sp}(x)$  and  $\mu_{sp}(y)$  when compared with  $\mu_{sp}(z)$ ] are (in D):

$$\mu_Q(z) = 2.5416 \sum_A Q_A z_A \quad (5)$$

$$\begin{aligned} \mu_{sp}(z) = -2.5416 \sum_A \left( \frac{5}{\sqrt{3}} \xi_A^{-1} P_{2z(A), 2p_z(A)} + \frac{7}{\sqrt{3}} \xi_A^{-1} P_{3z(A), 3p_z(A)} \right. \\ \left. + \frac{9}{\sqrt{3}} \xi_A^{-1} P_{4z(A), 4p_z(A)} \right). \end{aligned} \quad (6)$$

$$\begin{aligned} \mu_{pd}(x) = -2.5416 \sum_A \frac{7}{\sqrt{5}} \xi_A^{-1} \left( P_{3p_x(A), 3d_{x^2-y^2}(A)} + P_{3p_y(A), 3d_{xy}(A)} \right. \\ \left. + P_{3p_z(A), 3d_{xz}(A)} - \frac{1}{\sqrt{3}} P_{3p_x(A), 3d_{yz}(A)} \right) + \frac{9}{\sqrt{5}} \xi_A^{-1} \left( P_{4p_x(A), 4d_{x^2-y^2}(A)} \right. \\ \left. + P_{4p_y(A), 4d_{xy}(A)} + P_{4p_z(A), 4d_{xz}(A)} - \frac{1}{\sqrt{3}} P_{4p_x(A), 4d_{yz}(A)} \right). \end{aligned} \quad (7)$$

$$\begin{aligned} \mu_{pd}(y) = -2.5416 \sum_A \frac{7}{\sqrt{5}} \xi_A^{-1} \left( P_{3p_x(A), 3d_{xy}(A)} + P_{3p_z(A), 3d_{yz}(A)} \right. \\ \left. - P_{3p_y(A), 3d_{x^2-y^2}(A)} - \frac{1}{\sqrt{3}} P_{3p_y(A), 3d_{xz}(A)} \right) + \frac{9}{\sqrt{5}} \xi_A^{-1} \left( P_{4p_x(A), 4d_{xy}(A)} \right. \\ \left. + P_{4p_z(A), 4d_{yz}(A)} - P_{4p_y(A), 4d_{x^2-y^2}(A)} - \frac{1}{\sqrt{3}} P_{4p_y(A), 4d_{xz}(A)} \right). \end{aligned} \quad (8)$$

$$\begin{aligned} \mu_{pd}(z) = -2.5416 \sum_A \frac{7}{\sqrt{5}} \xi_A^{-1} \left( P_{3p_x(A), 3d_{xz}(A)} + P_{3p_y(A), 3d_{yz}(A)} \right. \\ \left. + \frac{2}{\sqrt{3}} P_{3p_z(A), 3d_{xz}(A)} \right) + \frac{9}{\sqrt{5}} \xi_A^{-1} \left( P_{4p_x(A), 4d_{xz}(A)} \right. \\ \left. + P_{4p_y(A), 4d_{yz}(A)} + \frac{2}{\sqrt{3}} P_{4p_z(A), 4d_{xz}(A)} \right). \end{aligned} \quad (9)$$

where  $Q_A = Z_A - P_{AA}$  is the net atomic charge on atom  $A$  and  $z_A$  its  $z$ -coordinate.

### 3. Numerical Results and Discussion

#### 3.1. Bond Lengths and Bond Angles

Table 1 lists the calculated and experimental bond lengths and angles of the hydrides and interhalogen compounds. Whereas bond lengths tend to differ a little bond angles are reproduced well. Calculations performed on larger molecules (invertomers of arsols [26] and conformers of germylethyl cations and anions [27]) are published elsewhere.

Table 1. Calculated and experimental bond lengths and angles

Compound	Bond lengths (Å)		Bond angles (degrees)	
	cal.	exp.	cal.	exp.
GeH <sub>4</sub>	1.726	1.527 [20]	109.5	109.5 [20]
AsH <sub>3</sub>	1.633	1.519 [21]	91.0	91.83 [21]
SeH <sub>2</sub>	1.551	1.460 [22]	92.1	91.0 [22]
BrH	1.464	1.414 [23]	—	—
BrF	1.618	1.7556 [24]	—	—
BrCl	1.810	2.138 [25]	—	—

Table 2. Calculated and measured values of the dipole moment.  $\mu_Q$ ,  $\mu_{sp}$ , and  $\mu_{pd}$  are the calculated contributions due to the net atomic charges and,  $sp$ - and  $pd$  atomic polarization, respectively

Compound [41]	Dipole moment (D)				exp.
	cal	cal. contributions			
		$\mu_Q$	$\mu_{sp}$	$\mu_{pd}$	
1-GeH <sub>4</sub>	0.18	-1.12	0.95	-0.01	---
ClGeH <sub>3</sub>	1.03	0.43	-0.46	1.92	2.124 [28]
AsH <sub>3</sub>	1.13	1.59	-3.40	2.93	0.22 [29]
AsF <sub>3</sub>	1.38	3.23	-3.15	1.30	2.57 [30]
AsCl <sub>3</sub>	0.15	1.95	-1.83	-0.27	1.6 [31]
As(CH <sub>3</sub> ) <sub>3</sub>	-2.10	0.34	-4.62	2.19	0.86 [32]
SeH <sub>2</sub>	0.02	0.02	-2.57	2.56	0.24 [22]
					0.62 [33]
Se(CH <sub>3</sub> ) <sub>2</sub>	1.05	0.51	-1.94	2.47	1.32 [34]
BrH	0.43	0.45	-1.72	1.73	0.788 [35]
BrF	1.45	1.48	-0.83	0.80	1.29 [24]
BrCl	0.34	0.26	-0.32	0.40	0.57 [25]
BrCH <sub>3</sub>	-1.83	-1.25	-1.71	1.13	1.797 [36]
BrCN	1.62	0.38	-0.57	1.80	2.94 [37]
BrNO [42]	-0.60	-1.50	0.66	1.30	1.80 [38]
BrC <sub>2</sub> H <sub>3</sub> [42]	-2.09	0.72	-2.32	0.45	1.417 [39]
BrC <sub>2</sub> H <sub>4</sub> [42]	-2.30	-1.47	-1.72	0.88	2.03 [40]

### 3.2. Dipole Moments

Table 2 presents the calculated dipole moments, the contributions  $\mu_Q$ ,  $\mu_{sp}$ , and  $\mu_{pd}$  and, where available, the measured moments. In all cases experimental geometries are used. The inclusion of  $d$  AOs is essential because of the large  $pd$  polarization contributions. Compared with the corresponding terms in molecules containing second-row atoms [6] both  $\mu_{sp}$  and  $\mu_{pd}$  are larger and more important for third-row molecules.

### 3.3. Ionization Potentials

As Table 3 shows, the calculated ionization potentials based on Koopmans' theorem [54] are systematically too high by 2–3 eV. This result is expected and consistent with previous findings for molecules containing first- and second-row atoms [55].

Table 3. Calculated and observed ionization potentials

Compound	Ionization potential (eV)	
	cal.	exp.
GeH <sub>4</sub>	15.319	12.3 [43]
FGeH <sub>3</sub>	15.472	15.0 [44]
ClGeH <sub>3</sub>	15.482	11.34 [44]
GeF <sub>4</sub>	19.210	16.03 [45]
AsH <sub>3</sub>	12.674	10.6 [46]
AsF <sub>3</sub>	14.446	—
AsCl <sub>3</sub>	13.366	11.7 [46]
As(CH <sub>3</sub> ) <sub>3</sub>	11.385	8.3 [46]
SeH <sub>2</sub>	12.239	9.88 [47]
Se(CH <sub>3</sub> ) <sub>2</sub>	11.328	8.2 [48]
BrH	13.904	11.62 [49]
BrF	14.098	11.9 [50]
BrCl	13.980	11.1 [50]
BrCH <sub>3</sub>	12.789	10.54 [51]
BrCN	13.431	11.95 [52]
BrNO	12.397	—
BrC <sub>2</sub> H <sub>3</sub>	11.793	9.82 [53]
BrC <sub>2</sub> H <sub>5</sub>	12.519	10.29 [49]

Nevertheless the calculated and the observed values both show the same trends. A good correlation between calculated and experimental ionization potentials in the series pyridine, phosphorin and arsenin [56] has been published elsewhere.

*Acknowledgements.* This work was supported by the Deutsche Forschungsgemeinschaft and by the Fonds der Chemischen Industrie. We thank Dr. I. L. Wilson for reading the manuscript.

### References

1. Dobosch, P.A.: Quantum Chemistry Program Exchange, Chemistry Department, Indiana University, Program: QCPE 141
2. Pople, J.A., Santry, D.P., Segal, G.A.: J. Chem. Phys. **43**, S 129 (1965)
3. Pople, J.A., Segal, G.A.: J. Chem. Phys. **43**, S 136 (1965)
4. Pople, J.A., Segal, G.A.: J. Chem. Phys. **44**, S 3289 (1966)
5. Pople, J.A., Beveridge, D.L.: Approximate molecular orbital theory. New York: McGraw-Hill 1970
6. Santry, D.P., Segal, G.A.: J. Chem. Phys. **47**, 158 (1967)
7. Santry, D.P.: J. Am. Chem. Soc. **90**, 3309 (1968)
8. Sabin, J.R., Santry, D.P., Weiss, K.: J. Am. Chem. Soc. **94**, 6651 (1972)
9. Sichel, J.M., Whitehead, M.A.: Theoret. chim. Acta (Berl.) **7**, 32 (1967)
10. Sichel, J.M., Whitehead, M.A.: Theoret. chim. Acta (Berl.) **11**, 220 (1968)
11. Sichel, J.M., Whitehead, M.A.: Theoret. chim. Acta (Berl.) **11**, 239 (1968)
12. Sichel, J.M., Whitehead, M.A.: Theoret. chim. Acta (Berl.) **11**, 254 (1968)
13. Sichel, J.M., Whitehead, M.A.: Theoret. chim. Acta (Berl.) **11**, 263 (1968)

14. Galasso, V., Trinajstić, N.: *Tetrahedron* **28**, 4419 (1972)
15. Slater, J. C.: *Phys. Rev.* **36**, 57 (1930)
16. Original STOs have  $n^*$  instead of  $n$  in the power of  $r$  (cf. 15). Since  $n^* = 3.7$  for third-row atoms the overlap integrals  $S_{\mu\nu}$  and repulsion integrals  $\gamma_{AB}$  (cf. Section 2.2 in the text) cannot be evaluated analytically. Therefore we use in accordance with [5]  $n$  ( $n = 4$ ) instead of  $n^*$
17. Cf. Eq. (B.20) in Ref. [5]
18. Mulliken, R. S.: *J. Chem. Phys.* **2**, 782 (1934)
19. Hinze, J., Jaffé, H. H.: *J. Phys. Chem.* **67**, 1501 (1963)
20. Lindemann, L. P., Wilson, M. K.: *J. Chem. Phys.* **22**, 1723 (1954)
21. Blevins, G. S., Jache, A. W., Gordy, W.: *Phys. Rev.* **97**, 684 (1955)
22. Jache, A. W., Moser, P. W., Gordy, W.: *J. Chem. Phys.* **25**, 209 (1956)
23. Naude, S. M., Verleger, H.: *Proc. Phys. Soc.* **63A**, 470 (1950)
24. Smith, D. F., Tidwell, M., Williams, D. V. P.: *Phys. Rev.* **77**, 420 (1950)
25. Smith, D. F., Tidwell, M., Williams, D. V. P.: *Phys. Rev.* **79**, 1007 (1950)
26. Hase, H. L., Schweig, A., Hahn, H., Radloff, J.: *Tetrahedron* **29**, 469 (1973)
27. Hase, H. L., Schweig, A.: *Tetrahedron* **29**, 1759 (1973)
28. Mays, J. M., Dailey, B. P.: *J. Chem. Phys.* **20**, 1695 (1952)
29. Loomis, C. C., Strandberg, M. W. P.: *Phys. Rev.* **81**, 798 (1951)
30. Buckingham, A. D., Raah, R. E.: *J. Chem. Soc.* **1961**, 5511
31. Grassi, U.: *Nuovo cimento* **14**, 461 (1937)
32. Lide, D. R.: *Spectrochim. Acta* **15**, 473 (1959)
33. Veselago, V. G.: *Optics and Spectroscopy* **6**, 286 (1959)
34. Chierici, I., Lombroso, H., Passerini, R.: *Compt. rend.* **237**, 611 (1953)
35. Compton, K. T., Zahn, C. T.: *Phys. Rev.* **23**, 781 (1924)
36. Shulman, R. G., Dailey, B. P., Townes, C. H.: *Phys. Rev.* **78**, 145 (1950)
37. Townes, C. H., Schawlow, A. L.: *Microwave Spectroscopy*. New York, McGraw-Hill 1955
38. Eagle, D. F., Weatherly, T. L., Williams, Q.: *J. Chem. Phys.* **30**, 603 (1959)
39. Huggill, J. A. C., Coop, I. F., Sutton, L. F.: *Trans. Faraday Soc.* **34**, 1518 (1938)
40. Groves, I. G., Sugden, S.: *J. Chem. Soc.* **1937**, 158
41. The sign of the dipole moments is defined as follows: The molecules are assumed to lie (cf. [42]) with their symmetry axis parallel to one Cartesian axis with the first atom in the formulae shown pointing in the positive direction of this axis
42. In these molecules the Br-N and Br-C bonds are assumed to lie parallel to one Cartesian axis with the Br atom pointing in the positive direction. Because of the lack of symmetry the three contributions  $\mu_Q$ ,  $\mu_p$  and  $\mu_{pd}$  have different directions in these cases
43. Neuert, H., Clasen, H.: *Z. Naturforsch.* **7A**, 411 (1952)
44. Craddock, S., Ishworth, E. A. V.: *Chem. Commun.* (1971) 57
45. Basset, P. J., Lloyd, D. R.: *J. Chem. Soc. A*, **1971**, 641
46. Cullen, W. R., Frost, D. C.: *Canad. J. Chem.* **40**, 390 (1962)
47. Price, W. C., Teegan, J. P., Walsh, A. D.: *Proc. Roy. Soc. (London)* **A201**, 600 (1950)
48. Magini, A., Trombetti, A., Zauli, C.: *J. Chem. Soc. B* **1967**, 153
49. Watanabe, K.: *J. Chem. Phys.* **26**, 542 (1957)
50. Irsa, A. P., Friedman, L.: *J. Inorg. Nucl. Chem.* **6**, 77 (1958)
51. Bralsford, R., Harris, P. V., Price, W. C.: *Proc. Roy. Soc. A* **258**, 459 (1960)
52. Herron, J. T., Dibeler, V. H.: *J. Am. Chem. Soc.* **82**, 1555 (1960)
53. Momigny, J.: *Bull. Soc. Chim. Belges*, **70**, 241 (1961)
54. Koopmans, T.: *Physica* **1**, 104 (1934)
55. Davis, D. W.: *Chem. Phys. Letters* **2**, 173 (1968)
56. Hase, H. L., Schweig, A., Hahn, H., Radloff, J.: *Tetrahedron* **29**, 475 (1973)

Prof. Dr. A. Schweig  
 Fachbereich Physikalische Chemie  
 Philipps-Universität Marburg  
 D-3550 Marburg (Lahn)  
 Biegenstr. 12  
 Federal Republic of Germany

## The Use of Site Symmetry in Constructing Symmetry Adapted Functions

R. L. Flurry, Jr.

Department of Chemistry, Louisiana State University in New Orleans,  
New Orleans, Louisiana 70122

Received March 5, 1973

It is shown that the application of a projection operator from a given group to a function is equivalent to the successive application of projection operators from factor groups of the starting group to that function. When used with the factor groups representing the site symmetry of a position and the simplest group of interchanges of positions, this concept provides a very simple method for obtaining symmetry adapted linear combinations of basis functions.

*Key words.* Symmetry adapted functions – Site symmetry – Projection operators – Factor groups

### Introduction

One of the most important uses of group theory by chemists lies in the construction of symmetry adapted linear combinations of basis functions. For most of the axial point groups (those with no more than one principal axis) the procedure involves a very simple explicit or implicit application of projection operators [1–3]. The procedure becomes more difficult for cubic and higher symmetries.

Frequently, symmetry adapted functions are constructed from subgroups of the point groups of the molecule under consideration (as, for example, in the use of the  $C_6$  group rather than  $D_{6h}$  for constructing the  $\pi$  molecular orbitals of benzene). In the axial point groups, the feasibility of such a process, and the appropriate choice of the subgroups is usually intuitively obvious to an experienced worker (and completely baffling to a neophyte). Again in the non-axial groups the process is much more difficult.

The purpose of this paper is to put the factorization process on a theoretically sound footing and to outline a simple systematic procedure for constructing symmetry adapted linear combinations (which, following Cotton [1], we will call SALC's) of the basis functions within any symmetry group. The method, which draws on the work of Altmann [4, 5], can be applied equally well to point groups of high symmetry (cubic, isosahedral, etc.), the symmetry groups of non-rigid molecules [6, 7], space groups, or any other problem where the group describing the system can be factored into appropriate subgroups.

The procedure will involve expressing the projection operators for the group under consideration as products of projection operators of two subgroups, the subgroup corresponding to the site symmetry of a basis function or set of basis functions and the subgroup corresponding to the simplest interchange of equi-

valent sites. Melvin [8] and Altmann [4, 5], have previously considered factorization of the projection operator; however, their factorizations are more restrictive than that presented here. The work of Melvin requires the concept of a "kernel" subgroup within a given irreducible representation; i.e., a subgroup onto which the irreducible representation under consideration maps as the totally symmetric irreducible representation. The factorization used by Altmann uses the concept of the poles of a rotation to derive an invariant subgroup of the group under consideration. The present work does not require either of these concepts.

### Groups as Products of Subgroups

Groups having more than one independent generator (*i.e.* more than one independent type of symmetry element) can be expressed as products of subgroups. (The product of two groups  $N$ , consisting of the set of elements  $\{N_i\}$  and  $C$ , consisting of the set of elements  $\{C_j\}$ , is the set of all products of elements  $\{N_i C_j\}$ .) The point group for any three-dimensional object can be expressed as the product of, at most, three independent subgroups. For systems with high symmetry, the factoring is usually not unique. The independent subgroups may themselves have more than one independent generator. Ultimately, however, the factoring can be continued until each subgroup has only one generator. These one-generator subgroups are cyclic in form, and consequently abelian.

In constructing a group as the product of two subgroups, two types of product are usually considered (see Altmann [4] for proofs); the direct product and the semidirect product. The direct product occurs when every element of one subgroup commutes with every element of the other. In semidirect products, the complete set of elements of one subgroup (the invariant subgroup) will commute with every element of the other subgroup. The notations are:

$$\text{Direct product} \quad G = N \times C \quad (1)$$

$$\text{Semidirect product} \quad G = N \wedge C \quad (2)$$

where in (2)  $N$  is the invariant subgroup. In the present work, we propose to construct a specific binary product structure for the point group and then to construct the SALC's in the full symmetry by building them up from the subgroups. The product, however, will not necessarily be constructed so that either group is invariant. It is, in fact, what is commonly called the "weak direct product". The direct product and semidirect product are special cases of this.

The proposed approach is to express the point group as the product of the site-symmetry group [9] of a given atom  $G_S$  and the simplest group which interchanges the considered atom with those equivalent to it,  $G_I$ .

$$G = G_I \cdot G_S \quad (3)$$

Note that a "dot" has been used to indicate the product since, depending upon the specific case, either a direct, a semidirect, or a weak direct product may be involved, and, if semidirect, either  $G_S$  or  $G_I$  may be invariant. Note also that the order of the group  $G_I$  is just equal to the number of equivalent atoms. The utility



of this factorization comes from the fact that in the simplest group of order  $n$  which interchanges  $n$  equivalent functions, i.e. functions which are identical except for interchange of origins (the group  $G_I$  in this case), each function generates a regular representation of the group [10]. Furthermore, every irreducible representation of a group is contained  $m_I$  times (where  $m_I$  is the dimension of the irreducible representation) in the regular representation of the group [11]. Thus, if every equivalent atom contained only one basis function, the irreducible representations of  $G_I$  would immediately yield the SALC's. This is the case, for example, for the  $\pi$  molecular orbitals of benzene.

Unfortunately, most cases of interest have more than one basis function centered on each atom. However, we shall show that all that is required to produce SALC's over the full group is to project independent functions within  $G_S$  and then combine these using the regular representation of  $G_I$ . The only "trick" is that all equivalent or degenerate basis functions must be generated by the operation of the elements of  $G_I$  on any one of the equivalent set.

### Projection Operators as Products of Projection Operators of Subgroups

Consider the form of the projection operator  $\hat{P}^{(\Gamma)}$  for the irreducible representation  $\Gamma$  in the group  $G$

$$\hat{P}^{(\Gamma)} = \sum_R \chi_R^{*(\Gamma)} \hat{R} \quad (4)$$

where  $\chi_R^{(\Gamma)}$  is the character of element  $R$  in representation  $\Gamma$  and  $\hat{R}$  is the corresponding group operation.  $\chi_R^{(\Gamma)}$  is defined as the trace of the matrix representing the operation  $R$  in the representation  $\Gamma$

$$\chi_R^{(\Gamma)} = \text{Tr } R^{(\Gamma)} = \sum_i R_{ii}^{(\Gamma)}. \quad (5)$$

Now if  $G$  is a product of  $G_I$  (having the operations  $\hat{S}$ ) and  $G_S$  (having the operations  $\hat{T}$ ) then

$$\hat{R} = \hat{S} \hat{T} \quad (6)$$

and  $R^{(\Gamma)}$ , the matrix representing  $\hat{R}$  in  $\Gamma$ , is the matrix direct product (outer or Kronecker product) of the matrices representing  $\hat{S}$  and  $\hat{T}$  in some representations, say  $\Gamma'$  and  $\Gamma''$  [12]

$$R^{(\Gamma)} = S^{(\Gamma')} \times T^{(\Gamma'')}. \quad (7)$$

From the properties of an outer product

$$\text{Tr } R^{(\Gamma)} = \text{Tr } S^{(\Gamma')} \text{Tr } T^{(\Gamma'')} \quad (8)$$

Thus,

$$\text{Tr } R^{(\Gamma)} = \sum_k \sum_l S_{kk}^{(\Gamma')} T_{ll}^{(\Gamma'')}. \quad (9)$$

Substituting Eqs. (6) and (9) into Eq. (4), and recognizing that the summation over  $R$  may be replaced by the double summation over  $S$  and  $T$ , we obtain

$$\hat{P}^{(\Gamma)} = \sum_S \sum_T \sum_k \sum_l S_{kk}^{*(\Gamma')} T_{ll}^{*(\Gamma'')} \hat{S} \hat{T} \quad (10)$$

Table 1. Site symmetry and interchange groups for some selected systems

Point group	Molecule	Ligand	$G_s$	$G_l$
$C_{3i}$	$AX_3$	$X$	$C_3$	$C_3$
	$AX_3Y$	$Y$	$C_{3v}$	$C_1$
$C_{4i}$	$AX_4$	$X$	$C_2$	$C_4$
	$AX_4YZ$	$Y(Z)$	$C_{4v}$	$C_1$
$C_{\infty}$	$AX_n$	$X$	$C_1$	$C_\infty$
$D_{2h}$	$AX_2Y_2Z_2$	$X(Y, Z)$	$C_{2v}$	$C_2$
$D_{2d}$	$AX_2Y_2$	$X(Y)$	$C_2$	$D_2$
$D_{3h}$	$AX_3$	$X$	$C_{2v}$	$C_3$
	$AX_3Y_2$	$Y$	$C_{3v}$	$C_2$
$D_{3d}$	$AX_6$	$X$	$C_2$	$S_6$
	$AX_4$	$X$	$C_{2v}$	$C_4$
$D_{4h}$	$AX_4Y_2$	$Y$	$C_{4v}$	$C_2$
	$X_6$	$X$	$C_{2v}$	$C_6$
$I_h$	$AX_4$	$X$	$C_{3v}$	$D_2$
$O_h$	$AX_6$	$X$	$C_{4v}$	$S_6$

which may be refactored to give

$$\begin{aligned}
 \hat{P}^{(I)} &= \sum_S \sum_k S_{kk}^{*(I')} \hat{S} \sum_T \sum_l T_{ll}^{*(I'')} \hat{T} \\
 &= \sum_S \chi_S^{*(I')} \hat{S} \sum_T \chi_T^{*(I'')} \hat{T} \\
 &= \hat{P}_{G_l}^{(I')} \hat{P}_{G_s}^{(I'')}
 \end{aligned} \tag{11}$$

Any SALC in  $G$  can be constructed by first projecting out the independent functions within the site-symmetry group  $G_s$  and then operating on this by the appropriate projection operator in the group of the simplest interchanges,  $G_l$ . However, we have already stated that, because of the generation of the regular representation, each irreducible representation of  $G_l$  occurs in a predetermined fashion. If degenerate functions in the starting basis set remain both degenerate and unique under the operations of  $G_s$ , the group  $G_l$  will have to be expanded to interchange all equivalent functions, and the site symmetry correspondingly reduced to a lower symmetry group. Such a situation occurs, for example, for  $p_x$  and  $p_y$  orbitals if a  $C_4$  site symmetry axis is present. The projection  $P^{(E)} p_x$  gives back pure  $p_x$  and similarly for  $p_y$ .

Once all degeneracies have been broken, the correlation theorem [13] and correlation tables can be used to immediately write down the desired SALC's.

### Applications

The systematic application of Eq. (11) for constructing SALC's will be outlined stepwise. As a specific example we will construct the molecular orbitals arising from the  $p$ -type atomic orbitals on the ligands in an  $AX_6$  octahedral complex.

Step 1: Factor the point group of the molecule into the product of  $G_l$  and  $G_s$ . In the octahedral complex ( $O_h$  point group) the ligand positions have  $C_{4v}$  site symmetry (four planes of symmetry and a  $C_4$  axis pass through each ligand). The

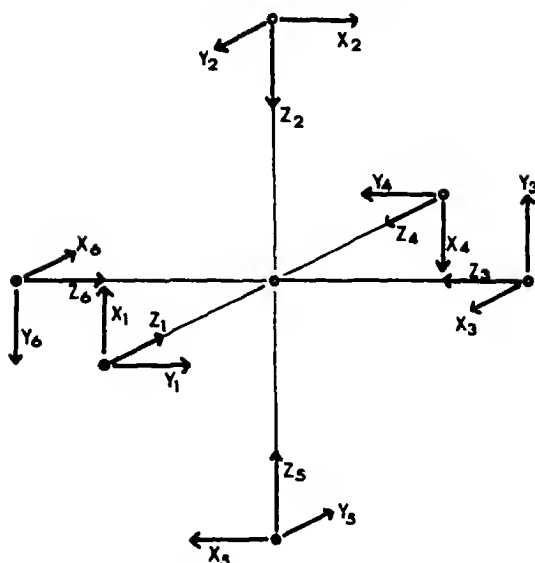


Fig. 1. Set of ligand  $p_x$ ,  $p_y$  and  $p_z$  orbitals in an octahedral complex

six ligands are interchanged by the operations of a group of order 6, the  $S_6$  group.  $G_S$  is simply the group of the symmetry elements which pass through the site.  $G_I$  may be found by factoring  $G$  into a product of abelian groups, by extracting  $G_S$  from this and by obtaining  $G_I$  from what is left. The usual product structure for  $O_h$  is [5]

$$O_h = D_2 \wedge D_3' \times C_i. \quad (12)$$

Factoring into abelian groups yields

$$O_h = C_2 \wedge C_2' \wedge C_3' \wedge C_2'' \times C_i. \quad (13)$$

The group  $C_{4v}$  is not immediately obvious from this; however,

$$C_{4v} \times C_i = D_{4h} = C_2 \wedge C_2' \wedge C_2'' \times C_i. \quad (14)$$

Thus

$$\begin{aligned} O_h &= C_3' \cdot C_i \cdot C_{4v} \\ &= S_6 \cdot C_{4v}. \end{aligned} \quad (15)$$

(Note that the invariance properties are lost when the  $C_2$  groups are commuted with the  $C_3$  group; consequently, the weak direct product must be used in Eq. (15).) Table 1 lists  $G_I$  and  $G_S$  for some selected structures.

Step 2: Orient the basis functions on each site so that equivalent functions are interchanged by the elements of  $G_I$ . If there are degeneracies in the basis functions which are not lifted by  $G_S$ , the orientations will not be unique. More will be said about this point later. A suitable choice for the  $p$ -type ligand atomic orbitals in the example is shown in Fig. 1.

Table 2. Correlation of the irreducible representations of  $O_h$  with those of the site-symmetry and interchange groups

$O_h$	$S_6$	$C_{4v}$	$D_{3d}$	$C_{2v}$
$A_{1g}$	$A_g$	$A_1$	$A_{1g}$	$A_1$
$A_{2g}$	$A_g$	$B_1$	$A_{2g}$	$B_1$
$E_g$	$E_g$	$A_1 + B_1$	$E_g$	$A_2 + B_2$
$T_{1g}$	$A_g + E_g$	$A_2 + E$	$A_{2g} + E_g$	$A_2 + B_1 + B_2$
$T_{2g}$	$A_g + E_g$	$B_2 + E$	$A_{1g} + E_g$	$A_1 + A_2 + B_2$
$A_{1u}$	$A_u$	$A_2$	$A_{1u}$	$A_2$
$A_{2u}$	$A_u$	$B_2$	$A_{2u}$	$B_2$
$E_u$	$E_u$	$A_2 + B_2$	$E_u$	$A_1 + B_1$
$T_{1u}$	$A_u + E_u$	$A_1 + E$	$A_{2u} + E_u$	$A_1 + B_1 + B_2$
$T_{2u}$	$A_u + E_u$	$B_1 + E$	$A_{1u} + E_u$	$A_1 + A_2 + B_1$

If degeneracies not lifted by  $G_S$  are associated with basis functions which are not mixed by the appropriate representations in  $G_S$ ,  $G_I$  must be expanded to interchange all equivalent functions. In the example, the  $p_x$  and  $p_y$  degeneracy is not removed by  $C_{4v}$ , nor are  $p_x$  and  $p_y$  mixed by  $\hat{P}^{(E)}$ . Thus, there are twelve equivalent basis functions which must be combined by  $G_I$ .  $G_I$  must be expanded to a group of order twelve. The group  $S_6 \cdot C_2'$ , where the axis in  $C_2'$  bisects adjacent pairs of ligands, has the proper order. If this  $C_2$  axis lies between ligands 1 and 2, it converts  $p_{x1}$  into  $p_{x2}$  and  $p_{y1}$  into  $-p_{y2}$ . The new  $G_I$  is  $D_{3d}$ . The new  $G_S, C_{2v}$ , is the group which reflects the site symmetry of the newly introduced  $C_2$  axis.

Step 3: Find the correlations of the irreducible representations of  $G_I$  and  $G_S$  with those of the complete group. For the example, these are shown in Table 2.

Step 4: Using the correlation theorem, find the irreducible representations of the SALC's which are expected to be derived from each type of basis function. The correlation theorem [13], which is, in fact, a simple application of the ideas of induced and subduced representations and the Frobenius reciprocity theorem, states that the irreducible representations in a group  $G$  spanned by a function, or set of functions, located at a site of lower symmetry  $G_S$  can be found by finding the irreducible representations spanned by the function in  $G_S$  and then correlating these with the irreducible representations of  $G$ . Each function having symmetry  $\Gamma'$  in  $G_S$  will contribute to every  $\Gamma$  in  $G$  which correlates with  $\Gamma'$ . From Table 2 it is seen that the  $A_1$  representation of  $C_{4v}$  correlates with the  $A_{1g}$ ,  $E_g$  and  $T_{1u}$  representations: thus the  $p_z$  ligand functions will appear in molecular orbitals of these symmetries. The  $E$  representation of  $C_{4v}$  correlates with the  $T_{1g}$ ,  $T_{2g}$ ,  $T_{1u}$  and  $T_{2u}$  representations of  $O_h$ . The  $p_x$  and  $p_y$  ligand orbitals will give rise to molecular orbitals of these symmetries.

Step 5: Find the independent basis functions within  $G_S$ . For the  $O_h$  system we are using, the  $p_z$  orbitals can be used directly in the  $S_6 \cdot C_{4v}$  factorization. They transform as the  $A_1$  representation within  $C_{4v}$ , and, being non-degenerate, are independent. The  $p_x$  and  $p_y$  orbitals transform as the  $E$  representation and are not

Table 3. Character tables for  $S_6$  and  $D_{3d}$  with correlations to  $O_h$ a)  $S_6$  (real form)

$O_h$	$S_6$	$E$	$2C_3$	$i$	$2S_6$
$A_{1g}$	$A_g$	1	1	1	1
$E_g$	$E_g$	2	-1	2	-1
	$A_u$	1	1	-1	-1
	$E_u$	2	-1	-2	1
$T_{1u}$	$A_u + E_u$	3	0	-3	0

b)  $D_{3d}$ 

$O_h$	$D_{3d}$	$E$	$2C_3$	$3C_2$	$i$	$2S_6$	$3\sigma_d$
	$A_{1g}$	1	1	1	1	1	1
	$A_{2g}$	1	1	-1	1	1	-1
	$E_g$	2	-1	0	2	-1	0
	$A_{1u}$	1	1	1	-1	-1	-1
	$A_{2u}$	1	1	-1	-1	-1	1
	$E_u$	2	-1	0	-2	1	0
$T_{1g}$	$A_{2g} + E_g$	3	0	-1	3	0	-1
$T_{2g}$	$A_{1g} + E_g$	3	0	1	3	0	1
$T_{1u}$	$A_{2u} + E_u$	3	0	-1	-3	0	1
$T_{2u}$	$A_{1u} + E_u$	3	0	1	-3	0	-1

independent. In the  $D_{3d} \cdot C_{2v}$  factorization, the projection operators from  $C_{2v}(G_S)$  must be used. We have, operating on say  $p_{x_1}$  and  $p_{y_1}$ ,

$$\begin{aligned}
 \hat{P}^{(A_1)} p_{x_1} &= p_{x_1} + p_{y_2} = \chi_{A_1} \\
 \hat{P}^{(A_2)} p_{y_1} &= p_{y_1} - p_{x_2} = \chi_{A_2} \\
 \hat{P}^{(B_1)} p_{x_1} &= p_{x_1} - p_{y_2} = \chi_{B_1} \\
 \hat{P}^{(B_2)} p_{y_1} &= p_{y_1} + p_{x_2} = \chi_{B_2}
 \end{aligned}
 \tag{16}$$

Step 6: Construct SALC's by using the projection operators from  $G_I$  and the functions from  $G_S$  obtained in step 5. The appropriate representations within  $G_I$  are obtained from the correlation tables obtained in step 3. The character tables for  $S_6$  and  $D_{3d}$  are given in Table 3, along with the representations leading to the desired representations in  $O_h$ . The results of applying the projection operators are (unnormalized):

For the  $p_z$  orbitals with  $G_I = S_6$

$$\begin{aligned}
 \phi(a_{1g}) &= \hat{P}^{(A_g)} p_{z_1} = p_{z_1} + p_{z_2} + p_{z_3} + p_{z_4} + p_{z_5} + p_{z_6} \\
 \phi(e_g) &= \hat{P}^{(E_g)} p_{z_1} = 2p_{z_1} - p_{z_3} - p_{z_5} + 2p_{z_4} - p_{z_5} - p_{z_6} \\
 \phi(t_{1u}) &= \hat{P}^{(A_u + E_u)} p_{z_1} = 3p_{z_1} - 3p_{z_4}
 \end{aligned}
 \tag{17}$$

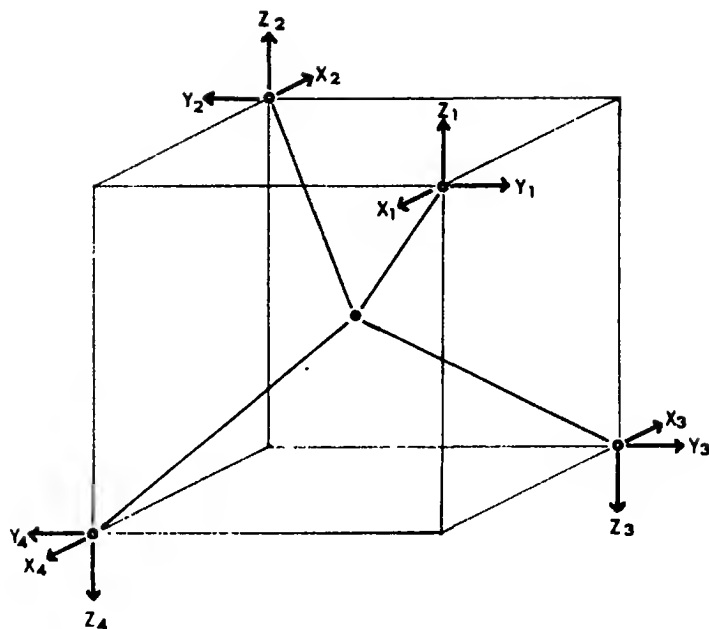


Fig. 2 Starting set of hydrogen centered basis vectors for the normal vibrations of methane

For the  $p_x$  and  $p_y$  orbitals with  $G_I = D_{3d}$  (the  $x_i$  and  $y_i$  represent the corresponding  $p_x$  and  $p_y$  orbitals)

$$\begin{aligned}
 \phi(t_{1g}) &= \hat{P}^{(A_{2g} + E_g)} \chi_{B_1} \\
 &= 2(x_1 - y_2) + y_1 - x_2 - x_3 + y_3 + 2(x_4 - y_5) + y_4 - x_5 - x_6 + y_6 \\
 \phi(t_{2g}) &= \hat{P}^{(A_{1g} + E_g)} \chi_{A_1} \\
 &= 4(x_1 + y_2) + y_1 + x_2 + x_3 + y_3 + 4(x_4 + y_5) + y_4 + x_5 + x_6 + y_6 \\
 \phi(t_{1u}) &= \hat{P}^{(A_{2u} + E_u)} \chi_{B_2} \\
 &= 4(y_1 + x_2) + x_1 + y_2 + x_3 + y_3 - 4(y_4 + x_5) - x_4 - y_5 - x_6 - y_6 \\
 \phi(t_{2u}) &= \hat{P}^{(A_{1u} + E_u)} \chi_{A_2} \\
 &= 2(y_1 - x_2) + x_1 - y_2 + x_3 - y_3 - 2(y_4 - x_5) - x_4 + y_5 - x_6 + y_6.
 \end{aligned} \tag{18}$$

These, of course, are not unique. In the present case, the use of the  $S_6$  subgroup of  $D_{3d}$  leads to a simpler, but equivalent set of functions.

$$\begin{aligned}
 \phi'(t_{1g}) &= \hat{P}^{(A_g + E_g)} \chi_{B_1} = 3(x_1 - y_2) + 3(x_4 - y_5) \\
 \phi'(t_{2g}) &= \hat{P}^{(A_g + E_g)} \chi_{A_1} = 3(x_1 + y_2) + 3(x_4 + y_5) \\
 \phi'(t_{1u}) &= \hat{P}^{(A_u + E_u)} \chi_{B_2} = 3(y_1 + x_2) - 3(y_4 + x_5) \\
 \phi'(t_{2u}) &= \hat{P}^{(A_u + E_u)} \chi_{A_2} = 3(y_1 - x_2) - 3(y_4 - x_5).
 \end{aligned} \tag{19}$$

These are in the form presented by Gray and Beach [14].

As a second example, consider the normal coordinates of methane. We will consider only the motions of the hydrogens relative to the carbon. The total

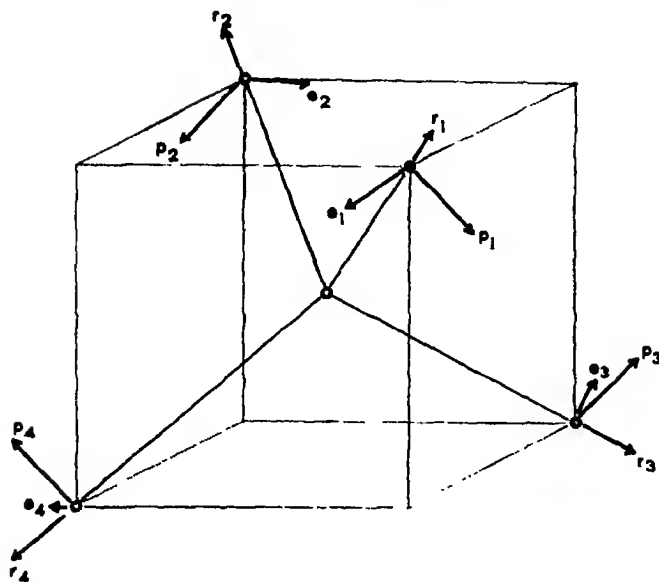


Fig. 3. Site-symmetry adapted hydrogen displacement vectors for methane

representation for a set of cartesian displacement vectors situated on the five atoms is  $A_1 + E + T_1 + 3T_2$  [9]. Of these, one  $T_2$  represents a translation of the entire molecule. This is the representation of the displacement vectors on the carbon; consequently, ignoring the carbon eliminates this from consideration. Of the remaining representations, which can all be obtained from the vectors on the four hydrogens, the  $T_1$  represents a rotation of the entire molecule. The vibrational representations are thus  $A_1 + E + 2T_2$ .

The steps are as follows:

1. The point group of the molecule is  $T_d$ . This can be factored into  $D_2 \wedge C_{3v}$ . The  $D_2$  is  $G_I$  and  $C_{3v}$  is  $G_S$ .
2. A suitable orientation for the displacement vectors is shown in Fig. 2. This orientation has been chosen so that the basis vectors match the site symmetry.
3. The correlations of  $D_2$  and  $C_{3v}$  with  $T_d$  are shown in Table 4.
4. From Table 4 it is seen that the  $A_1$  representation of  $C_{3v}$  correlates with the  $A_1$  and  $T_2$  representations of  $T_d$  while the  $E$  representation of  $C_{3v}$  correlates with the  $E$ ,  $T_1$  and  $T_2$  representations. The  $T_1$  representation is the representation of the rotations, therefore, we need not concern ourselves with it.

Table 4. Correlation of  $D_2$  and  $C_{3v}$  with  $T_d$ 

$T_d$	$D_2$	$C_{3v}$
$A_1$	$A$	$A_1$
$A_2$	$A$	$A_2$
$E$	$2A$	$E$
$T_1$	$B_1 + B_2 + B_3$	$A_2 + E$
$T_2$	$B_1 + B_2 + B_3$	$A_1 + E$

5. The independent basis vectors within  $C_{3v}$  are

$$\begin{aligned}\hat{P}^{(A_1)} X_1 &= X_1 + Y_1 + Z_1 = r_1 \\ \hat{P}^{(E)} X_1 &= 2X_1 - Y_1 - Z_1 = e_1 \\ \hat{P}^{(E)}(Y_1 - Z_1) &= Y_1 - Z_1 = p_1.\end{aligned}\quad (20)$$

These are shown in Fig. 3.

6. The final vibrational functions are (only one component of each degenerate function is given):

$$\phi(a_1) = \hat{P}^{(A_1)} r_1 = r_1 + r_2 + r_3 + r_4, \quad (21a)$$

$$\phi_1(t_2) = \hat{P}^{(B_1 + B_2 + B_3)} r_1 = 3r_1 - r_2 - r_3 - r_4, \quad (21b)$$

$$\phi(e) = \hat{P}^{(A_1)} e_1 = e_1 + e_2 + e_3 + e_4, \quad (21c)$$

$$\phi_2(t_2) = \hat{P}^{(B_1 + B_2 + B_3)} e_1 = 3e_1 - e_2 - e_3 - e_4. \quad (21d)$$

The functions represented by Eqs. (21a), (21b) and (21c) are completely equivalent to the functions  $v_1$ ,  $v_{3b}$  and  $v_{2a}$  given by Herzberg [15]. That of Eq. (21d) does not look like one of his forms. However, using the projection operator for only the  $B_3$  representation of  $D_2$  yields

$$\phi'_2(t_2) = \hat{P}^{(B_3)} e_1 = e_1 - e_2 - e_3 + e_4. \quad (22)$$

This is one of the  $v_4$  vibrations listed by Herzberg.

*Acknowledgement.* The author would like to thank Dr. T. H. Siddall, III and his students for calling attention to this problem.

## References

1. Cotton, F.A.: Chemical applications of group theory, 2nd ed., Chapter 6. New York: Wiley-Interscience 1971
2. Hall, L.H.: Group theory and symmetry in chemistry, Chapters 8 and 9. New York: McGraw-Hill Book Co., 1969
3. Leech, J.W., Newman, D.J.: How to use groups, Chapter 7. London: Methuen and Co., Ltd., 1969
4. Altmann, S.L.: Phil. Trans. Roy. Soc. **225** A, 216 (1963)
5. Altmann, S.L.: Rev. Mod. Phys. **35**, 641 (1963)
6. Longuet-Higgins, H.C.: Mol. Phys. **6**, 445 (1963)
7. Altmann, S.L.: Proc. Roy. Soc. A **298**, 184 (1967)
8. Melvin, M.A.: Rev. Mod. Phys. **28**, 18 (1956)
9. Flurry, Jr., R.L.: Int. J. Quantum Chem. **65**, 455 (1972)
10. Wilson, F.B., Decius, J.C., Cross, P.C.: Molecular vibrations, p. 113. New York: McGraw-Hill Book Co., 1955
11. Weyl, H.: The theory of groups and quantum mechanics, p. 306. Dover Publications, Inc., 1950
12. Ref. [10], p. 320
13. Ref. [10], p. 121
14. Gray, H.B., Beach, N.A.: J. Am. Chem. Soc. **85**, 2922 (1963)
15. Herzberg, G.: Infrared and raman spectra, p. 100. Princeton: D. Van Nostrand Co., Inc., 1945

Prof. Dr. R. L. Flurry  
Louisiana State University in New Orleans  
Department of Chemistry  
Lake Front  
New Orleans, Louisiana 70122  
USA



## An *ab initio* Study of the Hydration of Alkylammonium Groups

G. N. James Port and Alberte Pullman

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Theorique associe au C.N.R.S.  
13, rue P. et M. Curie, Paris 5e

Received May 9, 1973

*Ab initio* STO 3G calculations of the electronic structure and interaction energies with water of methyl and ethylammonium ions are reported. It is shown that the calculations predict a preferential attack at the  $\alpha$ -group (relative to the nitrogen), in agreement with experimental facts, and that successive ethylation reduces the favorable energy of hydration.

**Key words:** Hydration of alkylammonium

### Introduction

Alkylammonium groups occur frequently in pharmacologically active molecules; well-known examples are the cholinergic and adrenergic drugs. Generally, the nature of the alkyl substituents has a critical influence on the activity of the drug. For instance, in acetylcholine, replacement of the methyl groups of the cationic head by either ethyl groups or hydrogen atoms greatly reduces parasympathomimetic activity [1, 2]. In the adrenergic catecholamines, replacement of a hydrogen atom of the ammonium head by methyl and higher alkyl groups leads to a decrease in *alpha* activity but to an increase in *beta* activity [1, 3]. Similar modifications of the alkyl groups also produce changes in physico-chemical properties of alkylammonium ions in aqueous solution. These last changes have often been discussed in terms of the interactions of these ions with the water structure (see e.g. [4–6]).

This paper presents the results of a theoretical study of the electronic properties of methyl and ethyl ammonium ions and of their interactions with water. In particular we wish to examine whether there are significant changes in interaction with water when the methyl groups of a cationic head are replaced by ethyl groups. The calculations have been performed using the SCF *ab initio* method in an STO 3G basis [7]. The program Gaussian 70 [8] was used for the computations. Standard bond lengths were used ( $N-C = 1.49 \text{ \AA}$ ,  $C-C = 1.51 \text{ \AA}$ ,  $C-H = 1.09 \text{ \AA}$ ), and all angles were taken to be tetrahedral.

### Results

#### Atomic Charges of $N(CH_3)_4^+$ and $N(C_2H_5)_4^+$

There is some disagreement in the literature about the distribution of charge in alkylammonium groups. PCILO calculations by Pullman and Courrière [9] on acetylcholine, and CNDO calculations by George *et al.* [10] on alkylam-

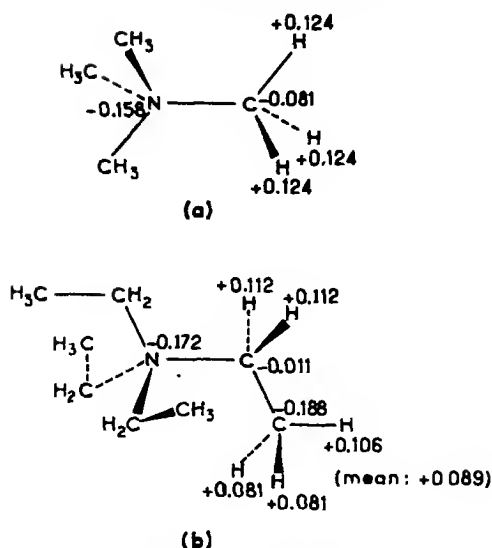


Fig. 1. Atomic charges in a) tetramethylammonium ion and b) tetraethylammonium ion. A positive charge corresponds to a deficiency of electrons

monium, show the positive charge distributed over the hydrogen atoms with the nitrogen and carbon atoms roughly neutral. This presents a picture of the alkylammonium head as a sphere of positive charge, with the nitrogen atom buried in the centre, and has obvious consequences for interaction with nucleophiles. INDO calculations on acetylcholine [11] predict on the contrary that the positive charge rests on the nitrogen atom and adjacent carbon atoms, which suggests different possibilities for attack by nucleophiles such as water.

The net atomic populations obtained by the STO 3G calculations for the tetramethyl and tetraethylammonium ions are shown in Fig. 1. In both ions the nitrogen atom is seen to be negative while the positive charge is located entirely on the hydrogen atoms. This last result clearly disagrees with the INDO predictions but is in accord with the PCIO and CNDO ones.

In the tetraethyl ion, which has more hydrogens, the positive charge is more spread out, but a striking fact is that it is not uniform. Thus the  $\alpha$  (methylene) hydrogens are more positive than the  $\beta$  (methyl) ones: +0.112 electron units as compared to +0.089 (mean value) respectively. Moreover, the  $\alpha$  carbons are less negative than the  $\beta$  ones, (-0.011 and -0.188 respectively), thus making the methylene groups *overall* considerably more positive than the terminal methyl groups or the nitrogen atom. These charge distributions, characteristic of ethyl and methyl groups, remain largely unaltered in the mixed ions, thus methyl has very similar atomic charges in  $\text{N}(\text{CH}_3)_4^+$  and in  $\text{NCH}_3(\text{C}_2\text{H}_5)_3^+$ .

In as far as hydration of alkylammonium ions is a predominantly long-range interaction depending on electrostatic forces, one might hope to predict the site of attack by water, or another nucleophile, from these charge distributions. Thus it seems most likely that such attack would occur at a position  $\alpha$  to the nitrogen rather than at this atom itself or at the hydrocarbon chain ends. To definitely

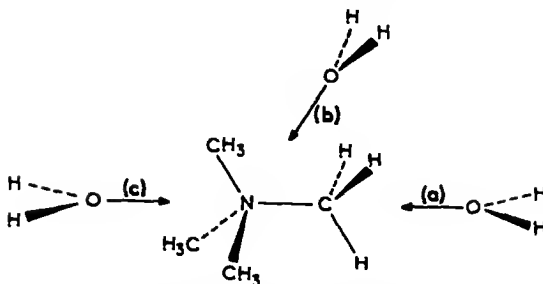


Fig. 2. Paths of approach of a water molecule to a tetramethylammonium ion. (a) direct approach towards a  $\text{CH}_3$  group along the prolongation of an  $\text{NC}$  bond. (b) bisecting approach along the bisectrix of two  $\text{NC}$  bonds. (c) axial approach on an  $\text{NC}$  axis [opposite to approach (a)]

establish the preferred hydration sites we have computed the interaction energy of a molecule of water with the alkylammonium ion in different orientations. These calculations are described below.

#### Hydration of the $\text{N}(\text{CH}_3)_4^+$ Ion

We have calculated the *ab initio* STO 3G energy of the system tetramethylammonium ion + water, as a function of intermolecular separation, for three possible paths of approach shown in Fig. 2.

The *direct* approach of water towards a methyl group along the  $\text{N}-\text{C}$  axis, (a) in Fig. 2, gives a minimum in total energy at  $\text{C}-\text{O}$  separation of  $3.1 \text{ \AA}$  and with an interaction energy (relative to the isolated molecules) of  $-5.4 \text{ kcal/mole}$ . An approach along a bisector of two  $\text{C}-\text{N}$  bonds, (b) in Fig. 2, shows a more favourable interaction ( $-6.4 \text{ kcal/mole}$  at  $\text{N}-\text{O}$  distance of  $3.7 \text{ \AA}$ ) possibly because the oxygen atom can interact simultaneously with two sets of positive  $\alpha$ -methyl hydrogens. Axial approach along a  $\text{C}-\text{N}$  axis as in (c) in Fig. 2 permits simultaneous interaction with three  $\alpha$  groups and is even more favourable. The minimum occurs at  $\text{N}-\text{O}$  separation of  $3.4 \text{ \AA}$  and with an interaction energy of  $-10.3 \text{ kcal}$ . The energy curve for this type of approach is shown in Fig. 3. It is interesting that the exact orientation of the water hydrogens is not found to be very important, leading to differences of less than  $0.5 \text{ kcal/mole}$  at the minimum.

The above results show a gain in stabilisation when the water molecule can interact with more than one  $\alpha$  group. The most favoured hydration site (Fig. 2(c)) permits simultaneous interaction with three  $\alpha$  groups, and there are four equivalent sites of this type. Although the sites where water interacts with only two  $\alpha$  groups are less favoured, there are six of them, and it is difficult to predict *a priori* which type of fixation will be more effective in the total hydration of  $\text{N}(\text{CH}_3)_4^+$ . We have accordingly added further water molecules in equivalent positions up to a maximum of four (which is the limit imposed by the program) and calculated the interaction energy for each complex, relative to the isolated components. These results are shown in Table 1 for both types of interaction and it is clear from the trends observed that even with six water molecules in the bisecting positions the total interaction energy will be less than for the complex with four water molecules in the most favoured axial sites.

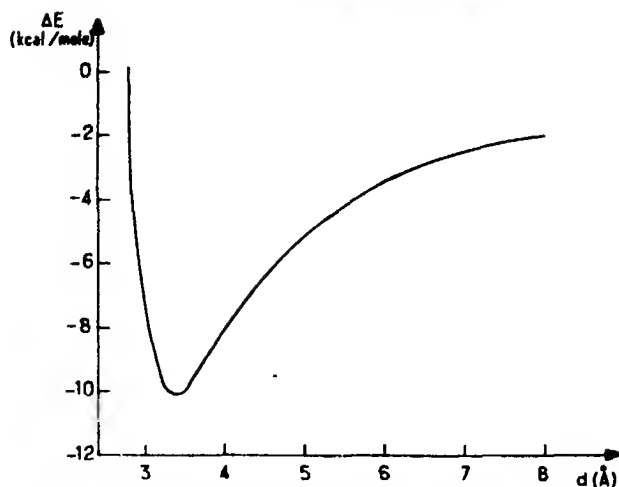


Fig. 3. The calculated stabilization energy for the interaction of tetramethylammonium ion with water as a function of intermolecular separation [water approaches as in Fig. 2, (c)].  $\Delta E$ : stabilisation energy;  $d$ : N—O distance

Table I. Calculated energies of interaction of tetramethylammonium ion with water molecules in two different types of site (STO 3G energies in kcal/mole)

Number of water molecules	Energy of interaction	
	(b) (bisectrix)	(c) (axial)
1	— 6.4	— 10.3
2	— 12.5	— 19.2
3	— 18.2	— 27.6
4	— 23.5	— 35.3
	...	
	...	

#### Hydration of the Ethyl-Substituted Ions

The single ethyl group in  $(\text{CH}_3)_3\text{N C}_2\text{H}_5^+$  offers the choice of  $\alpha$  or  $\beta$  attack by water as shown in Fig. 4. The  $\beta$  approach is exactly analogous to case (a) in  $\text{N}(\text{CH}_3)_4^+$  above but yields an interaction energy of (only)  $-3.4$  kcal/mole at a C—O distance of  $3.1$  Å. The  $\alpha$  approach (to the methylene group) is more favourable giving  $-5.6$  kcal/mole at a C—O distance of also  $3.1$  Å. Thus *within the ethyl group,  $\alpha$  attack by water is favoured over  $\beta$  attack*. Between different  $\alpha$  positions, an  $\alpha$  methylene group and an  $\alpha$  methyl group are about equally attractive to water. In connection with these results it is interesting to note that experimental observations of the crystal structures of a number of choline derivatives [12] show the halide counter-ion to be almost always located in the vicinity of an  $\alpha$  position on the alkyl substituents of the cationic head.

With regard to the total hydration of  $(\text{CH}_3)_3\text{N C}_2\text{H}_5^+$ , the ethyl group sterically hinders approach to one of the most favoured "axial" sites [Fig. 2(c)] but three

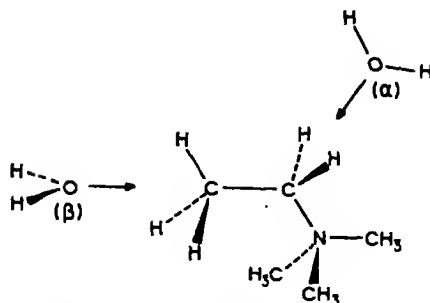


Fig. 4. The monoethyl substituted ion showing approach of water to  $\alpha$  and  $\beta$  positions

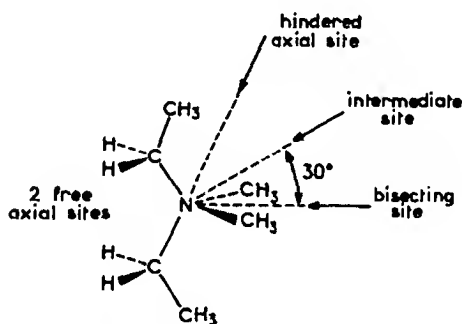


Fig. 5. Types of hydration site considered in the diethyl substituted ion

others are available. With these three occupied, the calculated energy of interaction is  $-26.7$  kcal/mole slightly less than for the first three water molecules in  $\text{N}(\text{CH}_3)_4^+$ . The limitations of the program do not permit the addition of one more water molecule, but it is clear that this could only enhance the difference relative to  $\text{N}(\text{CH}_3)_4^+$ .

In the diethyl-substituted ion  $(\text{CH}_3)_2\text{N}^+(\text{C}_2\text{H}_5)_2$  the two ethyl groups may be either *syn* or *anti*. In either case two of the "axial" sites are free and two are rather sterically hindered (see Fig. 5). The two water molecules in the hindered sites can occupy intermediate positions nearer to a bisecting type. Indeed the calculation shows a maximum interaction energy with these two molecules lying  $30^\circ$  above and below the plane of the N atom and the two methyl carbons (Fig. 5), with an N-O separation of  $3.6$  Å. The energies of interaction of one and two water molecules in these positions are given in Table 2 and indicate that the second ethyl substituent brings about slight further decreases relative to the tetramethylammonium ion.

In the triethylsubstituted ion,  $\text{CH}_3\text{N}^+(\text{C}_2\text{H}_5)_3$ , we have calculated the best energy of interaction with one water molecule. This molecule occupies an intermediate site of the type described above and the interaction energy shows again a further reduction (Table 2).

Table 2. Calculated hydration energies of methyl and ethylammonium ions as a function of number of water molecules and degree of ethylation (STO 3G energies in kcal/mole)

Number of water molecules	Energy of interaction			
	$\text{NMe}_4^+$	$\text{NMe}_3\text{Et}^+$	$\text{NMe}_2\text{Et}_2^+$	$\text{NMeEt}_3^+$
1	-10.3	-9.7	-8.7	-8.2
2	-19.2	-18.4	-16.0	.
3	-27.6	-26.7	..	—
4	-35.3	—	—	—

### Conclusions

The present calculations predict a gradual decrease in the magnitude of the hydration energy when the size and number of alkyl groups are increased. This trend is well illustrated by the figures of Table 2 and has two causes:

1) The charge distribution is smoothed out as the size of the alkyl chain increases, so that the positive charge located on the hydrogens becomes less concentrated.

2) Successive ethylation introduces steric hindrances, making the most favoured "axial" sites of  $\text{N}(\text{CH}_3)_4^+$  less and less available.

It is interesting to remark that the order of the trend obtained is parallel to the order of decreasing blocking activity exerted by quaternary ammonium derivatives at the neuro muscular junction [13].

*Note Added in Proof.* Further calculations show that due to the large positive charges on the hydrogens, direct C—H...O hydrogen bond formation is possible. Thus a water molecule forms such a hydrogen bond of -10 kcal/mole in  $\text{NMe}_4^+$ ; of -9 kcal/mole at an  $\alpha$  position in the ethylated ions and of -6 kcal/mole at a  $\beta$  position. The  $\alpha$  preference noted above is thus preserved. It seems unlikely that such interactions significantly change the overall hydration described above. Although twelve sites for H-bonding exist in  $\text{NMe}_4^+$ , they approach each other rather closely and after addition of four water molecules further water molecules introduce such intermolecular repulsions as to reduce rather than increase the total hydration energy which remains always slightly lower than that obtained with water in the axial type of site described above [14].

*Acknowledgments.* One of us (G.N.J.Port) acknowledges a fellowship from the Royal Society under its european programme.

### References

1. Ariëns, E.J., Simonis, A. M., Van Rossum, J.: Molecular pharmacology. Ariëns, E.J. ed., vol. 1, 169. New York: Acad. Press 1964.
2. Pullman, B., Courrière, Ph.: In press.
3. Kier, L. B.: Molecular orbital theory in drug research. New York: Acad. Press 1971
4. Kay, R. L., Evans, D. F.: J. Phys. Chem. **70**, 2325 (1966)
5. Gopal, R., Siddiqi: J. Phys. Chem. **72**, 1814 (1968)
6. Wen, W., Y., Hung, J. H.: J. Phys. Chem. **74**, 170 (1970)
7. Hehre, W. G., Stewart, R. F., Pople, J. A.: J. Chem. Phys. **51**, 2657 (1969)

8. Hehre, W. G., Lathan, W. A., Ditchfield, R., Newton, M. D., Pople, J. A.: Submitted to Q.C.P.E.
9. Pullman, B., Courrière, Ph., Coubeils, J. L.: *Mol. Pharm.* **7**, 397 (1971)
10. George, J. M., Kier, L. B., Hoyland, J. R.: *Mol. Pharm.* **7**, 328 (1971)
11. Beveridge, D. L., Radna, R. J.: *J. Am. Chem. Soc.* **93**, 3759 (1971)
12. Barrans, Y.: Thesis, University of Bordeaux (1971)
13. Ing, H. R.: *Progr. Drug. Res.*: **7**, 305 (1964)
14. Port, G. N. J., Pullman, A.: To be published

Prof. A. Pullman  
Institut de Biologie Physico-Chimique  
Fondation Edmond de Rothschild  
13, Rue Pierre et Marie Curie  
F-75005 Paris, France





# Direct Calculation of Ionization Potentials of Closed-Shell Atoms and Molecules

Lorenz S. Cederbaum\*.

Institut für Physikalische Chemie der Technischen Universität München

Received May 4, 1973

Vertical ionization potentials, electron affinities and information about quasi-particles can be obtained by using the technique of the single-particle propagator. The expansion of the self-energy part up to third order perturbation theory can be evaluated numerically, but does not lead, in most cases, to satisfying results. A theoretical and numerical analysis of the diagrammatic expansion of the self-energy part requires the introduction of a renormalized interaction and renormalized hole and particle lines.

**Key words:** Atomic and molecular structure · Electronic structure · Ionization · Many-body perturbation theory

## 1. Introduction

Many-body perturbation theories have been applied with success to nuclear structure [1], to high-density electron gas [2] and also to a wide range of other subjects as phonons, plasmons and superconductivity [3]. In addition, these theories have been also applied to atoms and molecules. In the case of atoms correlation energies [4, 5], dipole and quadrupole polarizabilities, shielding factors, transition probabilities [4], photodetachment cross sections [6], Fermi contact terms [7], open-shell SCF orbitals [29] and ionization potentials [5, 8] have been calculated. In the case of molecules the correlation problem [9], natural orbitals [30] and ionization potentials [10–12] have been treated (non semi-empirical calculations). The application of a many-body perturbation theory to ionization potentials is more than just an alternative for the usual calculation of these quantities, since Koopmans' defect [11]; the difference between the ionization potential obtained by Koopmans' theorem [13] and the exact one, is calculated directly without subtracting large numbers of nearly equal magnitude.

In previous discussions in [14] and applications to atoms [5, 8] and molecules [10] the self-energy part has been expanded up to second order. Extensive calculations, however, have shown that, at least for small molecules, the expansion of the self-energy part up to second order is far from being able to reproduce the experimental results [11, 12, 15]. Therefore, a more elaborate form of perturbation theory has to be derived which is done in the following sections. Quantitative results are discussed for the nitrogen molecule. Further calculations for  $F_2$ ,  $C_2H_2$ ,  $H_2O$  and  $H_2CO$  are given in [15].

\* Present Address: Physik-Department, Technische Universität München, Deutschland.

## 2. General Theory

In this section the general theory of Green's one-particle functions and of the self-energy part is discussed which is essential for the following sections.

In the following text we use the occupation number formalism in which the operators  $a_i$  and  $a_i^\dagger$  are the annihilation and the creation operators for a particle in the state  $|i\rangle$  satisfying the anti-commutation relations

$$[a_i, a_j^\dagger]_+ = \delta_{ij}, \quad [a_i, a_j]_+ = [a_i^\dagger, a_j^\dagger]_+ = 0, \quad (2.1)$$

$$i \equiv (i, \sigma_i).$$

Let a system of  $N$  interacting fermions in the ground state be described by  $\Psi_0^N$ . In the Heisenberg representation the Green's one-particle function is defined by

$$G_{kk'}(t, t') = -i \langle \Psi_0^N | T \{ a_k(t) a_{k'}^\dagger(t') \} | \Psi_0^N \rangle, \quad (2.2)$$

$$a_k(t) = e^{iHt} a_k e^{-iHt}, \quad a_k \equiv a_k(0),$$

$T$  = Wick time-ordering operator.

The spectral representation can be written as

$$G_{kk'}(\omega) = \int_{-\infty}^{\infty} G_{kk'}(t, t') e^{i\omega(t-t')} d(t-t') \quad (2.3)$$

$$= \sum_l \frac{\langle \Psi_0^N | a_k | \Psi_l^{N+1} \rangle \langle \Psi_l^{N+1} | a_{k'}^\dagger | \Psi_0^N \rangle}{\omega + A_l + i\eta} + \sum_l \frac{\langle \Psi_0^N | a_{k'}^\dagger | \Psi_l^{N-1} \rangle \langle \Psi_l^{N-1} | a_k | \Psi_0^N \rangle}{\omega + I_l - i\eta}$$

$$\eta \rightarrow 0^+$$

where the relations

$$I_l = E_l^{(N-1)} - E_0^{(N)}$$

$$A_l = E_0^{(N)} - E_l^{(N+1)}$$

hold<sup>1</sup>.  $E_0^{(N)}$  is the ground state energy of the  $N$ -electron system and  $E_l^{(N-1)}$  and  $E_l^{(N+1)}$  the energy of a  $l$ -th state of the  $(N-1)$ - and  $(N+1)$ -electron systems, respectively. Hence, the problem of calculating ionization potentials and electron affinities is equivalent to the problem of calculating the poles of the one-particle Green's functions. Putting the energies corresponding to the molecular geometry of the initial state for  $E_l^{(N-1)}$  and  $E_l^{(N+1)}$  the vertical ionization potentials (VIPs) and electron affinities (VEAs), respectively, are obtained. In order to compare them with the Franck-Condon maxima of the experimental ionization spectrum, they should be corrected for vibrational effects [16].

The Green's one-particle function can be expanded in the time representation by means of a perturbation theory which is described in detail in [17-19]. It is therefore not considered here. It should be noted that, starting out from the equations of motion for the Green's  $n$ -particle functions, an infinite coupled set of equations results [20]. The uncoupling of this set of equations yields the same result as the perturbation expansion of the Green's functions.

The perturbation series of the Green's function, however, does not have a form which allows the evaluation of its poles. On the other hand the Dyson equation in

<sup>1</sup> In order to simplify, the first ionization potential is denoted by  $I_0$ .

the  $\omega$ -representation

$$\begin{aligned} G &= G^0 + G^0 \Sigma G, \\ G^{-1} &= G^{0-1} - \Sigma \end{aligned} \quad (2.4)$$

is especially appropriate for this purpose. The poles of  $G$  correspond to the zeros of the eigenvalues of  $G^{-1}$ . In Eq. (2.4) the quantity  $\Sigma_{sm}$  is the self-energy part and  $G_{kk}^0$  is the free Green's function

$$\begin{aligned} G_{kk}^0(t, t') &= -i \langle \Phi_0 | T \{ a_k(t) a_k^\dagger(t') \} | \Phi_0 \rangle \\ &= \delta_{kk} e^{-i\epsilon_k(t-t')} \begin{cases} -i & t > t', \quad k \notin \{\text{occ}\} \\ i & t \leq t', \quad k \in \{\text{occ}\} \end{cases} \end{aligned} \quad (2.5)$$

where  $\Phi_0$  is the ground state eigenfunction of  $H_0$  and  $\{\text{occ}\}$  is the set of the orbitals occupied in this state.

For the above reason  $\Sigma$  shall be considered in more detail:

Let the Hamiltonian be given by

$$\begin{aligned} H &= H_0 + H_w, \\ H_0 &= \sum h_i a_i^\dagger a_i, \end{aligned} \quad (2.6)$$

$$H_w = \frac{1}{2} \sum V_{ijkl} a_i^\dagger a_j^\dagger a_l a_k$$

where

$$\begin{aligned} h_i &= \langle \varphi_i(1) | h(1) | \varphi_i(1) \rangle, \\ V_{ijkl} &= \langle \varphi_i(1) \varphi_j(2) | V(1, 2) | \varphi_k(1) \varphi_l(2) \rangle \end{aligned} \quad (2.7)$$

are the matrix elements of the one-particle operator  $h$  and the two-particle operator  $V$ , respectively, and  $\{\varphi_i\}$  is a complete set of one-particle wave functions. From the perturbation expansion of  $G$  and the Dyson equation we obtain the expansion of  $\Sigma$ . The perturbation series of the Green's function and hence the expansion of the self-energy part is greatly simplified by the application of a well-known diagrammatic method. The following definitions hold:

$$\begin{aligned} \begin{array}{c} i \quad k \\ | \quad | \\ t \quad t' \\ | \quad | \\ i' \quad k' \end{array} &\equiv i G_{kk}^0(t, t') \\ \begin{array}{c} i \quad j \\ \diagup \quad \diagdown \\ \bullet \\ \diagdown \quad \diagup \\ k \quad l \end{array} &\equiv -i V_{ij[kl]} = -i(V_{ijkl} - V_{ijlk}) \\ \begin{array}{c} i \quad j \\ \diagup \quad \diagdown \\ \text{---} \\ \diagdown \quad \diagup \\ k \quad l \end{array} &\equiv -i V_{ijkl} \end{aligned} \quad (2.8)$$

The rules to draw the terms of a certain  $n$ -th order of the diagrammatic expansion of an element of  $\Sigma$  are:

( $\alpha 1$ ) The elements of each Graph are  $n$   $V_{ij[kl]}$ -points and  $(2n-1)$   $G^0$ -lines. The elements of one kind can be connected only with elements of the other kind.

( $\alpha 2$ ) Graphs, which split into two graphs by removing a single  $G^0$ -line do not belong to  $\Sigma$  according to the Dyson equation.

( $\alpha 3$ ) All topologically not equivalent linked graphs with two free indices have to be drawn according to the rules ( $\alpha 1$ ) and ( $\alpha 2$ ).

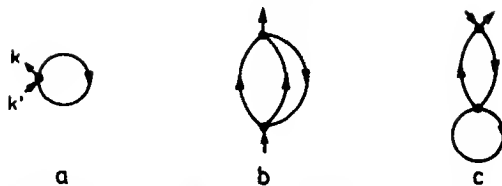


Fig. 1. The graphs of the first (a) and second (b, c) order of the expansion of  $\Sigma$

In Fig. 1 the first and second order of the expansion are reported as an example.

The graphs which are obtained via  $(\alpha 1)$ – $(\alpha 3)$  by replacing the  $V_{ij[kl]}$ -points by the wiggle are called Feynman graphs. Any of the graphs treated here contains several Feynman graphs, e.g.

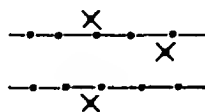
$$\equiv \quad (2.9)$$

Except for the rules  $(\beta 6)$  and  $(\beta 7)$ , all the following rules for evaluation of diagrams are also valid for Feynman diagrams.

$(\beta 1)$  Join the free indices  $k, k'$  of an  $n$ -th order diagram of the expansion of  $\Sigma_{kk'}(t, t')$  with a  $e^{-i\omega(t'-t)}$ -line, which shall be denoted by

$$e^{-i\omega(t'-t)} \equiv \begin{array}{c} t' \\ | \\ t \end{array}$$

$(\beta 2)$  Draw the  $(n-1)$  horizontal lines  $\cdots$  between successive pairs of  $V$ -points according to:



Any part of the diagram between two successive  $V$ -points is called a block.

$(\beta 3)$  Each  $G^0$ -line and  $e^{-i\omega(t'-t)}$ -line cut by a horizontal line supplies an additive contribution to the denominator of the block, namely:

yields the dominator  
 $+ \omega - h_i - h_j + h_l + \dots, \quad i, j \notin \{\text{occ}\}, l \in \{\text{occ}\}.$

$(\beta 4)$  Multiply the interactions  $V_{ij[kl]}$ , the contributions of the blocks and a factor  $(-1)^{\Sigma_i + \Sigma_s}$ ; then sum over the internal indices ( $\Sigma_i$  is the number of hole lines,  $\Sigma_s$  is the number of loops).

$(\beta 5)$  Each of the  $n!$  time ordered diagrams has to be evaluated separately.

$(\beta 6)$  Each graphs has to be multiplied by  $2^{-q}$  where  $q$  is the number of permutations of two  $G^0$ -lines in the diagram leaving the diagram unchanged (identity transformation).

( $\beta 7$ ) The sign of the  $V_{ij[kl]}$ -points is not uniquely determined. The proper sign of the graph follows from a comparison with the sign of a Feynman graph which is contained in it.

As an example the terms for the graph of the second order of the expansion of  $\Sigma_{kk'}(\omega)$  are obtained with the help of these rules.

$$\text{graph}(b) = \frac{1}{2} \sum_{\sigma_2, \sigma_3} \frac{V_{ks[f]} V_{k's[f]}}{\omega + h_s - h_f - h_i}; \quad (2.10)$$

$$\sigma_2 = \{s \in \{\text{occ}\}; f, l \notin \{\text{occ}\}\}, \quad \sigma_3 = \{s \notin \{\text{occ}\}; f, l \in \{\text{occ}\}\}.$$

In the above perturbation theory the Hamiltonian of Eq. (2.6) has been used. Analogous results can also be obtained by choosing the perturbation  $H_w$  in a different way. In this work we choose now the Hartree-Fock operator

$$H_{\text{HF}} = \sum \varepsilon_i a_i^\dagger a_i, \quad (2.11)$$

$$H_w = \frac{1}{2} \sum V_{ijkl} a_i^\dagger a_j^\dagger a_l a_k - \sum \left( \sum_{l \in \{\text{occ}\}} V_{ilij} \right) a_i^\dagger a_j$$

as the unperturbed operator  $H_0$  and the canonical HF-orbitals as the one-particle wave functions  $\phi_i$ . The  $\varepsilon_i$  are here the HF-energies. Very important reasons for this choice are the availability of HF calculations and that the HF-energies of occupied orbitals of closed-shell systems provide relatively good approximations for the VIPs. The absence of bound excited (unoccupied) HF states, as might be the case for neutral atoms, leads to a slow convergence of the perturbation expansion [21]. The convergence of the expansion for correlation energies of atoms was greatly improved by using the  $V(N-1)$  potential of Kelly [4]. For closed-shell molecules, however, with many bound excited HF states we have no reason to assume a better convergence of the expansion of the self-energy part especially in calculating VIPs.

A consequent use of the  $V(N-1)$  potential leads to a great expense for systems with many electrons since these excited states should be calculated in the potential field of  $N-1$  other electrons [22] and therefore this potential is different for different excited states.

If the HF operator is taken to be  $H_0$  then on account of the first term of  $H_w$ , graphs containing  $G_{kk}^0(t, t+0)$ -lines need no more be considered [18]. This means that only  $h_i$  has to be replaced by  $\varepsilon_i$  in the expressions for the graphs and that, in addition, the number of graphs shrinks considerably. Graph (b) is the only remaining graph in Fig. 1, hence Koopmans' theorem is obtained in the first order of the expansion. As already mentioned in the introduction the expansion of the self-energy part up to the second order is in many cases far from being able to reproduce the experimental results. Therefore we have to use a more elaborate form of perturbation treatment. A general way to do so is to renormalize quantities as the interaction, vertex and particle- and hole-lines. An example for a renormalized interaction is given by the random phase approximation (see Section 4.1). By the renormalization of hole- and particle-lines a transformation of  $G^0$ -lines into  $G$ -lines is understood. A graph, which is to be "dressed", is often referred to as "skeleton". By self-consistent perturbation theory we mean the self-consistent

solution of the equations

$$\Sigma = \text{diagram 1} + \text{diagram 2} + \dots$$

$$\text{diagram 3} = \text{diagram 4} + \text{diagram 5}$$
(2.12)

where the symbol

$$\text{diagram 6} \equiv i G_{kk'}(t, t')$$
(2.13)

is used.

In all renormalizations attention must be paid to overcounting.

In the next section the pole strengths of the Green's functions are discussed and in Section 4 a self-energy part which is appropriate for calculating VIPs is derived. We use the nitrogen molecule as an example throughout this manuscript. Additional numerical calculations are given in [12, 15, 23].

### 3. Restriction to Low Energy; Main and Secondary Poles

In the following text some properties of the self-energy part which are essential for this work are mentioned. It should be noted that the self-energy part  $\Sigma$  also has poles. The following notation will be used: the poles of graph (b) in Eq. (2.10) with the index set  $\sigma_2$  are denoted by  $\Sigma_{+1}, \Sigma_{+2}, \dots$  in the order of increasing energy and the poles with the index set  $\sigma_3$  are denoted by  $\Sigma_{-1}, \Sigma_{-2}, \dots$  in the order of decreasing energy. The energy intervals  $(\Sigma_l, \Sigma_{l+1})$ ,  $l \neq -1$ , will be denoted by the index  $l$  for  $l \geq 1$ , by  $l+1$  for  $l \leq -2$  and the interval  $(\Sigma_{-1}, \Sigma_{+1})$  by  $l=0$ . These definitions are illustrated in Fig. 2, where a schematic plot of  $\Sigma_{ii}(\omega)$  is given.

In the picture of the Hartree-Fock quasiparticles the poles  $\Sigma_{-k}$  for  $k \geq 1$  correspond to processes in which one particle is separated and simultaneously another one is excited to an unoccupied orbital. In a similar manner the poles  $\Sigma_k$  for  $k \geq 1$  are connected with processes in which one particle is added and another one excited. It can be shown that exactly one pole of the eigenvalue  $D_{ii}(\omega)$  of the Green's function  $G$  is situated between any two successive poles of the self-energy

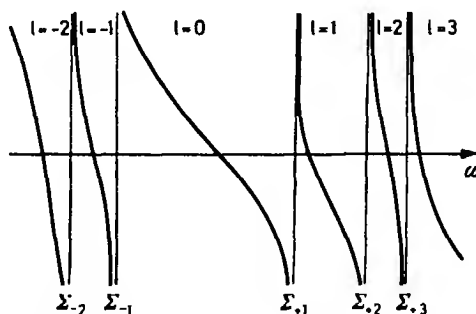


Fig. 2. A schematic plot of  $\Sigma_{ii}$  for the definition of its poles

part [12, 24]. This statement is easily verified for the second order of the perturbation expansion with the help of Eq. (2.10) [12]. An essential consequence of this statement is, that the zeros of the eigenvalue  $D_{kk}^{-1}$  of  $G^{-1}(\omega)$  which do not satisfy the conditions

$$\omega_k \in (\Sigma_{-1}, \Sigma_{+1}), \quad |\omega_k - \Sigma_{\pm 1}| \gg 0 \quad (3.1)$$

have to be treated in a different way than the zeros which do satisfy these conditions. The present work deals only with such solutions  $k$  of the Dyson equation which satisfy the above conditions. The remaining VIPs and VEAs are treated elsewhere [12, 24]. *A common theory for all VIPs and VEAs is developed in [24] but it leads to a great numerical expense.*

A perturbation expansion of the Green's one-particle function gives information not only about the energy values of the VIPs and VEAs but also about the utility of the one-particle picture for ionization and particle addition processes. The desired information is provided by the pole strengths of the Green's function, which are considered in the following. First of all, some definitions:  $P_{kl}$  is the pole strength of the  $k$ -th eigenvalue  $D_k(\omega)$  of the Green's function matrix  $G$  to the pole in the interval  $l$  and  $a_{ijkl}$  is the pole strength of  $G_{ij}(\omega)$  to the same pole ( $k, l$ ).

$$P_{kl} = \lim_{\omega \rightarrow \omega_{kl}} D_k(\omega) \cdot (\omega - \omega_{kl}), \quad (3.2)$$

$$a_{ijkl} = \lim_{\omega \rightarrow \omega_{kl}} G_{ij}(\omega) \cdot (\omega - \omega_{kl}),$$

$\omega_{kl}$  is the energy-coordinate of the pole and  $S_{ikl} = S_{ik}(\omega_{kl})$  is an element of the eigenvector matrix  $S(\omega)$  to the same pole.

$$D^{-1} = S^+ G^{-1} S. \quad (3.3)$$

By means of the Dyson equation one obtains directly for the pole strength

$$P_{kl}^{-1} = 1 - \left. \frac{\partial}{\partial \omega} (\Sigma + \varepsilon)_k \right|_{\omega_{kl}} \quad (3.4)$$

where  $A_k$  means  $(S^+ A S)_{kk}$  generally.

In order to interpret the pole strengths one has to start out from their definition. By use of Eq. (2.3) one gets

$$a_{ikl} = P_{kl} |S_{ikl}|^2 = |\langle \psi_{kl}^{N-1} | a_l | \psi_0^N \rangle|^2, \quad (3.5)$$

$$k \in \{\text{occ}\}, l = 0, -1, -2, \dots,$$

$$k \notin \{\text{occ}\}, l = -1, -2, \dots,$$

$$1 \geq a_{ikl} \geq 0.$$

$|\psi_{kl}^{N+1}\rangle$  and  $|\psi_{kl}^{N-1}\rangle$  are true states of the  $(N+1)$ -particle system and  $(N-1)$ -particle system, respectively.

Hence,  $a_{ikl}$  is the probability to find the state  $a_l |\psi_0^N\rangle$  in the state  $|\psi_{kl}^{N-1}\rangle$ , that means the projection of a true state on a fictive one. The result of "removing" a HF-particle from the true ground state is described by  $a_l |\psi_0^N\rangle$ , the true state after ionization is described by  $|\psi_{kl}^{N-1}\rangle$ . In case  $a_{ikl}$  is not much less than 1, the "removing" is nearly identical with the ionization.

Analogous conclusions are obtained for the pole strengths  $a_{ikl}$ , where

$$k \in \{\text{occ}\}, l = 1, 2, 3, \dots,$$

$$k \notin \{\text{occ}\}, l = 0, 1, \dots$$

One has only to replace  $a_i$  by  $a_i^+$ ,  $\psi_{kl}^{N-1}$  by  $\psi_{kl}^{N+1}$  and "removing" by "addition". For the example of  $N_2$  we calculated the following pole strengths:

$$P_{2\sigma_u,0} = 0.87,$$

$$P_{3\sigma_g,0} = 0.91,$$

$$P_{1\pi_u,0} = 0.94.$$

In order to compare those pole strengths of  $D$  with those of  $G$  we must know the values of  $S_{ikl}$ .

All numerical results indicate that the eigenvectors of  $G$  associated with eigenvalues  $\omega_{kl}$  which have pole strengths,  $P_{kl}$ , not much less than 1, may be approximated very well by unit vectors. This is mainly due to the fact that the inequality

$$|(e_i + \Sigma_{ii}) - (e_j + \Sigma_{jj})| \gg |\Sigma_{ij}|, \quad i \neq j \quad (3.6)$$

is satisfied for all our examples.

The eigenvectors of  $G$  were calculated for the example of  $N_2$  with the self-energy part in second order and are compiled in Table 1.

If the ionization is well described by the "removing" of a quasi-particle in the state  $k$ , then the corresponding Green's function in the range of the solution must show a similar form as the free function  $G_{kk}^0(\omega)$ . The self-energy part can be decomposed into imaginary and real part for real values of  $\omega$  according to

$$\Sigma = \Sigma^R + \Sigma^I. \quad (3.7)$$

Table 1. Eigenvectors of  $G^{-1}$  at  $\omega = \omega_{k0}$  calculated with the self-energy part in second order for  $N_2$ . The basis-set used is described in [11]

Species	$k$							
	$2\sigma_g$	$2\sigma_u$	$1\pi_u$	$3\sigma_g$	$1\pi_g$	$3\sigma_u$	$4\sigma_g$	$2\pi_u$
$2\sigma_g$	0.9932	0	0	0.0346	0	0	0.0011	0
$2\sigma_u$	0	0.9995	0	0	0	-0.0031	0	0
$1\pi_u$	0	0	1.0000	0	0	0	0	0.0329
$3\sigma_g$	-0.1146	0	0	0.9990	0	0	0.0069	0
$1\pi_g$	0	0	0	0	1	0	0	0
$3\sigma_u$	0	-0.0121	0	0	0	0.9999	0	0
$4\sigma_g$	-0.0025	0	0	0.0050	0	0	0.9980	0
$2\pi_u$	0	0	0.0038	0	0	0	0	0.9995
$5\sigma_g$	0.0224	0	0	0.0265	0	0	0.0631	0
$4\sigma_u$	0	0.0294	0	0	0	-0.0146	0	0



Starting out from the Dyson equation and expanding  $(\epsilon + \Sigma^R)_k$  about  $\omega_{kl} = (\epsilon + \Sigma^R)_k|_{\omega_{kl}}$  to the linear term and expanding  $\Sigma_k^I$  about  $\omega_{kl}$  to the constant term one obtains

$$D_k(\omega \approx \omega_{kl}) = \frac{P_{kl}}{\omega - \omega_{kl} + i\tau_{kl}^{-1}} + f_{kl}(\omega \approx \omega_{kl}) \quad (3.8)$$

$$\tau_{kl}^{-1} = P_{kl} \cdot \Sigma_k^I(\omega_{kl}) \cdot \begin{cases} -1, & -\omega_{kl} = \text{VIP} \\ +1, & -\omega_{kl} = \text{VEA} \end{cases}$$

where  $f_{kl}(\omega)$  is a correction function for the higher orders of the expansion about  $\omega_{kl}$ . This expansion leads to the above result only if  $S$  is independent of  $\Sigma^I$  for  $\eta \rightarrow 0^+$ . But this is true, because  $S$  has no poles in the open interval between two successive poles of  $\Sigma$ . Hence, if  $f_{kl}(\omega \approx \omega_{kl})$  is unessential and  $P_{kl} \approx 1$ , then  $D_k(\omega \approx \omega_{kl})$  has the form of a quasi-particle propagator. The corresponding quasiparticle has a finite life time  $\tau_{kl}$ .

Equation (3.5) and the important relation

$$\sum_{k,l} a_{ijkl} = \delta_{ij}, \quad (3.9)$$

which is proved by removing the unit operators between the Heisenberg operators  $a_k$  and  $a_k^+$ , in Eq. (2.3), imply, that, if  $D_k$  takes the form of a quasi-particle propagator, there is only one dominant pole strength,  $P_{kl}$  for each eigenvalue  $D_k$  of  $G$ . This pole shall be referred to as main pole, the remaining poles shall be called secondary poles. The properties of the secondary poles are treated extensively elsewhere [12, 24].

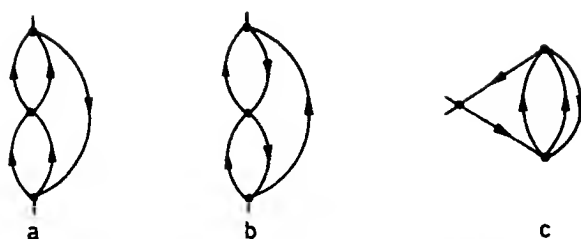
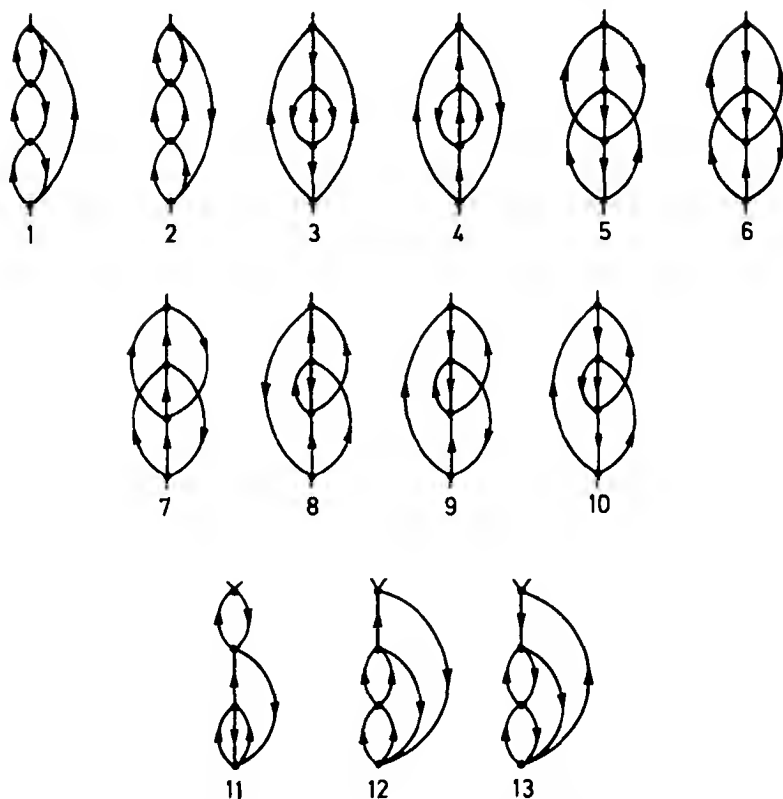
#### 4. Inclusion of Higher Orders

The expression for the graph of the second order of the expansion of the self-energy part has already been given as an example in Section 2. All numerical calculations for Ne, N<sub>2</sub>, F<sub>2</sub>, H<sub>2</sub>O, C<sub>2</sub>H<sub>2</sub> and CO<sub>2</sub> [11, 12, 15, 25] show uniquely that the expansion of the self-energy part to the second order does not at all suffice to evaluate the VIPs via the Dyson equation. As an example the VIPs of N<sub>2</sub> calculated with the self-energy part of second order are compiled in Table 2.

For the above one cannot expect that taking in account the third order will be sufficient. An evaluation of additional orders turns out to be completely useless, as the number and the magnitude of the expressions go up explosively with the

Table 2. Results for the three lowest ionizations of N<sub>2</sub>. R2 = VIP calculated with self-energy part in second order. RF = VIP calculated with the present theory. All energies in eV

Species	Exp. VIP [28]	R2	RF
2σ <sub>u</sub>	18.78	17.00	18.59
1π <sub>u</sub>	16.98	16.96	16.83
3σ <sub>g</sub>	15.60	14.44	15.50

Fig. 3. The graphs of third order of the expansion of  $\Sigma$ Fig. 4. The graphs of fourth order of the expansion of  $\Sigma$ 

order of the expansion. The graph of the second order contains already 4 terms with 3 indices each. The graphs of the third order reported in Fig. 3 contain 84 terms with 5 indices each. One notes, however, that these graphs can still be calculated numerically. The necessary numerical methods are described elsewhere [12, 15]. The graphs of the fourth order, reported in Fig. 4, contain already 3120 expressions with 7 indices each. Therefore it is not possible to evaluate the total fourth order numerically.

One has to find out which graphs yield the important contributions, and then to calculate or, at least, to estimate these to infinite order.

As a first step towards the solution of this problem it is reasonable to investigate the essential results of the solid-state physics and nuclear physics with respect to their application in atomic and molecular physics.

#### 4.1. A Short Comparison with Electron Gas and Nuclear Matter

It is common in the treatment of metals to introduce a scaling parameter  $r_s$ , describing the volume per electron in an electron gas. It can be shown that, in the high-density case ( $r_s \rightarrow 0$ ), the sum of the ring diagrams ("RPA")

$$\Sigma \approx \text{ring diagram 1} + \text{ring diagram 2} + \text{ring diagram 3} + \dots \quad (4.1)$$

describes the self-energy part sufficiently well [26, 19]. With  $r_1$  denoting the Bohr radius the volume per electron in the electron gas is

$$\Omega_e = \frac{4\pi}{3} (r_s \cdot r_1)^3. \quad (4.2)$$

Let  $R_n$  be the radius of the  $n$ -th Bohr orbit and  $N_{\text{eff}}$  the number of electrons considered, then from

$$R_n \approx n^2 \cdot r_1 \cdot Z_{\text{eff}}^{-1} \quad (4.3)$$

it follows that:

$$r_{s_n} \approx n^2 Z_{\text{eff}}^{-1} N_{\text{eff}}^{-1/3}. \quad (4.4)$$

In regard of the inner electrons with  $Z_{\text{eff}} \gg 1$  the relation  $r_s \ll 1$  follows from Eq. (4.4). Thus, it is reasonable to evaluate the correlation energy for such atoms by means of a model of a high-density non-uniform electron gas. Considering only the outer-shell electrons on the other side, one obtains, e.g. for phosphorus  $r_s \approx 1.3$ . Therefore, the approximation of the self-energy part by ring diagrams does not work well in evaluating VIPs of atoms and especially not of molecules. It should be mentioned here that Brueckner [27] already tried to evaluate correlation energies of atoms with the help of a model of a high-density non-uniform electron gas. He obtained the important result that it has no sense to consider the atom as an uniform high-density electron gas and that the density gradient does not converge in the case of inhomogeneity.

For the treatment of nuclear matter, however, circumstances are different. As before, it is also here common to introduce a parameter similar to  $r_s$ . With a standing for the effective range of the interaction and  $r_0$  standing for the average distance between the interacting particles one speaks of low-density, if  $r_0/a \gg 1$ . Galitzki [26] showed that each hole line in a graph of the self-energy part is associated with a factor  $a/r_0$ . The sum of the graphs containing only one hole line

which are called ladder graphs, dominates in the low density case.

$$\begin{aligned}\Sigma &\approx \text{[rectangle with semi-circle on right]} + \text{[square with diagonal line]} \\ \text{[square]} &= \text{[wavy line]} + \text{[ladder graph with wavy top line]} \end{aligned} \quad (4.5)$$

Theoretical as well as numerical comparisons of

$$\begin{array}{cc} \text{[two vertices with two arcs]} & \text{[two vertices with four arcs]} \end{array} \quad (4.6)$$

and

$$\begin{array}{cc} \text{[two vertices with two arcs]} & \text{[two vertices with four arcs]} \end{array} \quad (4.7)$$

show, however, that, for atoms and molecules, the contribution of the first mentioned graphs are comparable with those of the second mentioned graphs and, as a rule, even exceed them [12].

All these considerations lead to the result that atoms and molecules are located in the region between intermediate density and high-density, which means in a region, where it cannot be expected that just a few graphs dominate. This renders a reasonable evaluation of the self-energy part much more difficult.

#### 4.2. Antigrahs and Renormalized Interaction

All graphs of third order and a number of graphs of fourth order have been evaluated for the systems Ne, N<sub>2</sub>, F<sub>2</sub>, H<sub>2</sub>O, CO<sub>2</sub>, C<sub>2</sub>H<sub>2</sub> and H<sub>2</sub>CO. The contributions of some of these graphs were of the same order of magnitude as the contributions of the graphs of second order. If we assume the convergence of the perturbation expansion, some of the graphs must compensate each other at least partly. This is confirmed by numerical results. The assumption that the contributions of the single graphs go down with increasing order of the expansion cannot be maintained. Hence, it is useless to consider as many graphs of a certain order of the expansion as possible without closer examination. It should be examined whether there exists a small "parameter" which gives us an idea how to select the important terms in the expansion of  $\Sigma$ . In order to do so first of all some of the specific features of graphs of third order are considered in more detail. Each graph is split into its 6 time orders where the nomenclature given in Fig. 5 will be used.

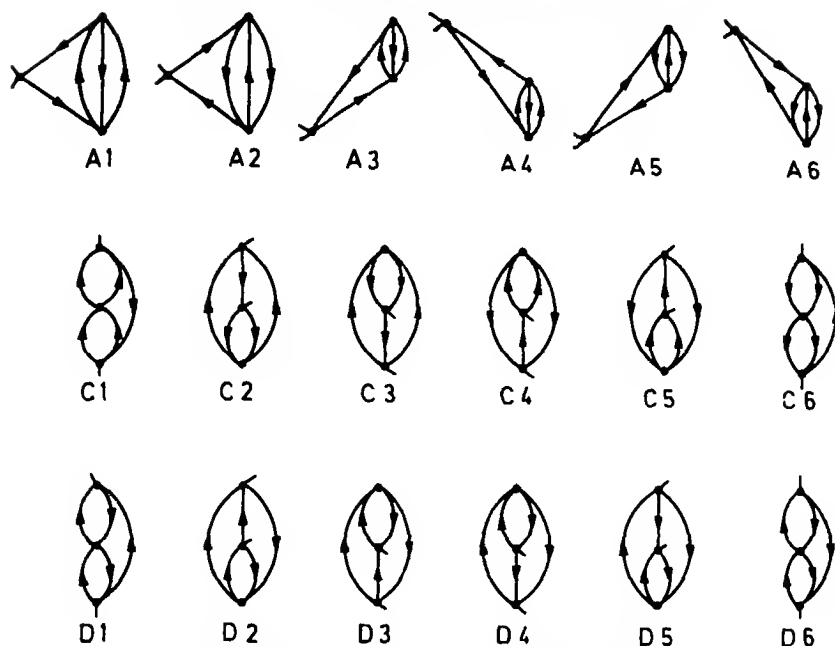


Fig. 5. The notation of the time ordered graphs of third order

In the following only diagonal elements of the self-energy part of third and higher orders are considered (see Section 3). It is easily shown, that, for  $i = j$  the relations

$$\begin{aligned} A_3 &= A_4, A_5 = A_6, \\ X_2 &= X_3, X_4 = X_5, X = C, D \end{aligned} \quad (4.8)$$

hold. There remain 12 different time ordered graphs.

It is rather difficult to determine additional *exact* relations, if it is possible at all. Expressing the terms of the graphs  $C_1 - C_6$  and  $A_1 - A_6$  in an appropriate way facilitates relations as, e.g.

$$\begin{aligned} |C_1| &\ll |C_6| & \omega \approx \epsilon_k, k \in \{\text{occ}\} \\ A_1, C_6 &< 0 & \text{for all } k. \\ A_2, C_1 &> 0 \end{aligned} \quad (4.9)$$

As one can see from the explicit expressions, it is difficult to build up relations for  $D_1 - D_6$  and between  $D$  and  $C$  graphs, as the  $D$ -graphs contain 8 expressions each, which cannot be collected as easily as for the  $C$ -graphs. Only simple relations like

$$D_1 < 0, D_6 > 0 \quad \omega \approx \epsilon_k \quad (4.10)$$

may be established [12].

In order to determine additional relations among the graphs we make the approximation that the most important contributions to  $\Sigma$  arise from one occupied

Table 3. The contributions of the graphs of third order for  $N_2$  [eV]<sup>a</sup>

Orb.	A1	A2	2.A3	2.A5	C1	D1	2.C2	2.D2	2.C4	2.D4	C6	D6
$2\sigma_u$	-1.78	1.73	0.41	-0.39	0.10	-0.13	0.29	-0.25	-0.10	-1.79	-2.35	2.10
$3\sigma_u$	-1.70	1.77	0.42	-0.44	0.15	-0.16	0.35	-0.25	-0.12	-1.23	-1.56	1.34
$1\pi_u$	-1.80	1.77	0.44	-0.41	0.32	-0.30	0.77	-0.64	-0.08	-0.11	-0.61	0.65
$1\pi_g$	-1.76	1.39	0.38	-0.39	0.27	-0.46	0.03	0.16	-0.63	0.66	-0.40	0.32
$3\sigma_g$	1.16	1.10	0.11	-0.11	0.21	-0.19	0.03	0.33	-0.03	0.04	-0.03	0.03
$4\sigma_g$	1.35	1.27	0.17	-0.17	0.29	-0.27	0.02	0.36	-0.03	0.04	-0.02	0.02
$2\pi_u$	-1.52	1.40	0.22	-0.23	0.54	-0.68	0.18	0.07	-0.05	0.07	-0.04	0.04
$5\sigma_g$	-1.70	1.52	0.31	-0.29	1.04	-1.20	0.09	0.66	-0.12	0.11	-0.08	0.07
$\Delta_{-1} = -37.90$ eV, $\Sigma_{+1} = 24.86$ eV.												

<sup>a</sup> Each graph has been evaluated at  $\omega = \text{main pole}$ .

orbital and one unoccupied orbital of suitable symmetry. With this approximation we deduce relations which are confirmed by a large number of numerical results. As well *ab initio* calculations for the above mentioned systems as semi-empirical HF-calculations for a larger number of molecules [25, 23] were available. The graphs of third order have been calculated in all these cases. As an example, the data of these graphs for the nitrogen molecule are compiled in Table 3. The numerical and theoretical results indicate that the following relations are valid:

$$0 \geq A1 \approx -A2, 0 \leq C1 \approx -D1 \quad (4.11)$$

$$0 \geq C6 \approx -D6, A3 \approx -A5,$$

$$C2 \approx -D2 \quad k \in \{\text{occ}\}$$

$$C4 \approx -D4 \quad k \notin \{\text{occ}\}$$

for


$$\omega \approx \omega_{k0}, \quad |\omega - \Sigma_{\pm 1}| \gg 0.$$

Hence, there exist pairs of graphs distinguishing themselves by nearly compensating each other. Such graphs shall be referred to as antigraphs. Although they do not completely compensate each other, both together have to be considered as one quantity in the perturbation theory. In third order there exist 5 different pairs of antigraphs.

With  $\chi_j^{(n)}$  and  $g_k^{(n)}$  standing for a pair of antigraphs and an arbitrary graph of  $n$ -th order which is no antigraph, respectively, the self-energy part can be written:

$$\Sigma_{ii}(\omega) = \sum_{j,n} [\chi_j^{(n)}]_{ii} + \sum_{k,n} [g_k^{(n)}]_{ii}. \quad (4.12)$$

Both time ordered diagrams of the self-energy part of second order may be considered as antigraphs, too, since


(4.13)

$\omega \approx \omega_{k0}$

hold.

This pair of antigraphs of second order,  $\chi_1^{(2)}$ , is essential for most orbitals and

$$|[\chi_1^{(2)}]_{ii}| \gg |[\chi_s^{(3)}]_{ii}|, \quad s = 1, 2, 3, 4, 5 \quad (4.14)$$

is valid for these orbitals, which indicates that the first sum converges quickly and can be estimated quite well by summing up to third order. Therefore, if the graphs of third order are calculated explicitly, then only graphs  $\{g_k^{(n)}\}$ , for  $n \geq 4$ , have to be considered. We try to take account of the essential contributions of these graphs by introducing a time dependent effective interaction and renormalized hole- and particle-lines. The renormalized interaction shall be symbolized by

$$\text{[Diagram: A square with four external lines, two on the left and two on the right, representing a renormalized interaction]} = -i V_{\text{eff}}(t, t'). \quad (4.15)$$

This renormalized interaction contains all graphs beginning with two free indices at the time  $t'$  and ending with two free indices at the time  $t$ :

$$\text{[Diagram: A square with four external lines, two on the left and two on the right, representing a renormalized interaction]} = \text{[Diagram: A cross with four external lines]} + \text{[Diagram: A circle with two external lines]} + \frac{1}{2} \text{[Diagram: A circle with two external lines]} + \frac{1}{4} \text{[Diagram: A circle with two external lines]} + \dots \quad (4.16)$$

In this equation the factor of rule ( $\beta 6$ ) in Section 2 is explicitly put in front of the graph.

It is evident, that the renormalized interaction satisfies the recurrence formula

$$\begin{aligned} \text{[Diagram: A square with four external lines, two on the left and two on the right, representing a renormalized interaction]}^{(n)} &= \text{[Diagram: A cross with four external lines]} + \text{[Diagram: A circle with two external lines]} + \frac{1}{2} \text{[Diagram: A circle with two external lines]} + \dots \text{[Diagram: A square with four external lines, two on the left and two on the right, representing a renormalized interaction]}^{(n-1)} \\ \text{[Diagram: A square with four external lines, two on the left and two on the right, representing a renormalized interaction]}^{(0)} &= \text{[Diagram: A cross with four external lines]} \end{aligned} \quad (4.17)$$

Numerical results prove that the calculation of  $V_{\text{eff}}$  by use of Eq. (4.17) converges slowly. It is useless, for this reason, just to evaluate a few orders and it is, therefore, necessary to estimate the rest.

Substituting the interaction by the renormalized one and paying attention to repetitions of graphs one gets in second order

$$\frac{1}{2} \text{[Diagram: A circle with two external lines]} \rightarrow \frac{1}{2} \text{[Diagram: A circle with two external lines]} = \frac{1}{2} \text{[Diagram: A circle with two external lines]} = \frac{1}{2} \text{[Diagram: A circle with two external lines]} + \text{[Diagram: A circle with two external lines]} + \frac{1}{4} \text{[Diagram: A circle with two external lines]} + \dots \quad (4.18)$$

All  $\omega$ -dependent graphs of third order as well as the graphs (1) and (2) in Fig. 4 are contained in the renormalized graph of second order.

Analogously

$$\frac{1}{2} \text{[Diagram: A circle with two external lines]} \rightarrow \frac{1}{2} \text{[Diagram: A circle with two external lines]} = \frac{1}{2} \text{[Diagram: A circle with two external lines]} + \text{[Diagram: A circle with two external lines]} + \frac{1}{2} \text{[Diagram: A circle with two external lines]} + \dots \quad (4.19)$$

By use of the antigraph hypothesis one obtains

$$\begin{aligned}
 & \frac{1}{2} \left( \text{graph}_1 + \text{graph}_2 \right) \underset{i \in \{\text{occ}\}}{\approx} \sum_{n=2}^3 \sum_{S_n=1}^{1,4} [\chi_{S_n}^{(n)}]_{ii} + \text{graph}_3 + \text{graph}_4 \\
 & + \frac{1}{4} \left[ \text{graph}_5 + \text{graph}_6 \right] + \text{graph}_7 + \text{graph}_8 + \text{graph}_9 + \quad (4.20) \\
 & + \text{graph}_{10} + \text{graph}_{11} + \frac{1}{8} \left[ \text{graph}_{12} + \text{graph}_{13} + \text{graph}_{14} + \text{graph}_{15} + \text{graph}_{16} \right] + \dots
 \end{aligned}$$

for the renormalized graph. It means, that graphs which arise from the graphs of third order which are not antigraphs are considered.

For  $i \notin \{\text{occ}\}$  an analogous expression results [12]. It must be observed, however, that the error of the estimation arises in the fourth order, because *all  $\omega$ -dependent graphs of third order are considered explicitly* in Eq. (4.20).

All drawn graphs on the right hand side of Eq. (4.20) are of the type  $\theta_{k_n}^{(n)}$ . These graphs have been numerically evaluated for all above examples and provide extremely large contributions [12, 15].

In the following we want to try to evaluate the renormalized graph of second order.

First, we have to make use of the fact that the following series are equal:

$$\begin{aligned}
 & \left[ \text{graph}_1 \right]_{ii} + \left[ \text{graph}_2 + \text{graph}_3 \right]_{ii} + \left[ \text{graph}_4 \right]_{ii} + \dots \\
 & \left[ \text{graph}_5 \right]_{ii} + \left[ \text{graph}_6 + \text{graph}_7 \right]_{ii} + \left[ \text{graph}_8 \right]_{ii} + \dots \quad (4.21)
 \end{aligned}$$

analogously

$$\begin{aligned}
 & \left[ \text{graph}_9 \right]_{ii} + \left[ \text{graph}_{10} + \text{graph}_{11} \right]_{ii} + \left[ \text{graph}_{12} \right]_{ii} + \dots \\
 & \left[ \text{graph}_{13} \right]_{ii} + \left[ \text{graph}_{14} + \text{graph}_{15} \right]_{ii} + \left[ \text{graph}_{16} \right]_{ii} + \dots \quad (4.21a)
 \end{aligned}$$



In crude approximation these series may be estimated by geometric series. The quotient of the first two terms of the series in Eq. (4.21) and the series in Eq. (4.21a) shall be referred to as  $q_l^{(a)}$  and  $q_l^{(b)}$ , respectively.

$$q_l^{(a)} = \left[ \text{graph} \right]_{11}^{-1} \cdot \left[ \text{graph} + \text{graph} \right]_{11} \quad (4.22)$$

$$q_l^{(b)} = \left[ \text{graph} \right]_{11}^{-1} \cdot \left[ \text{graph} + \text{graph} \right]_{11} \quad (4.22a)$$

In order to estimate the renormalized graph in Eq. (4.20) the graphs

$$\text{graph} \quad \text{graph} \quad (4.23)$$

are left, which give only small contributions to the examples considered. But this must be checked for each new example.

For the example of  $N_2$  the quotients  $q^{(a)}$  and  $q^{(b)}$  are:

$$\begin{aligned} q_{2\sigma_u}^{(a)} &= -0.89, & q_{1\pi_u}^{(a)} &= -0.99, & q_{3\sigma_g}^{(a)} &= -0.87, \\ q_{2\sigma_u}^{(b)} &= -0.30, & q_{1\pi_u}^{(b)} &= -0.50, & q_{3\sigma_g}^{(b)} &= -0.33. \end{aligned}$$

This means that, as  $C4$  yields only a small contribution, the contribution of the effective interaction is approximately given by the terms of second order and the half of  $2 \cdot D4$ .

The graph of second order is the first graph of the expansion, which depends on  $\omega$ . The constant graph  $C$  in Fig. 3 is the first graph which does not depend on  $\omega$ . Renormalizing this graph according to Eq. (4.19) one obtains in fourth order the  $\omega$ -dependent graphs

$$\text{graph} \quad \text{graph} \quad (4.24)$$

These graphs have a special meaning which is explained in the next section.

The graphs in Eq. (4.19) which do not depend on  $\omega$ , are not of importance unless the corresponding antigraph relations in Eq. (4.11) are not satisfied sufficiently (Appendix A).

### 4.3. Self-Consistent Perturbation Theory

In Section 4.2 a renormalization of the interaction has been described. Numerical calculations show that the obtained results for VIPs are largely improved by

considering the renormalization. For further improvement we also consider the renormalization of hole- and particle-lines. The renormalization of the "skeleton"  $\times \bigcirc$  is reduced to the HF-problem by use of Eq. (2.11) with

$$\Sigma = \text{diagram 1} - \text{diagram 2} \quad (4.25)$$

If we choose but the self-energy part of second order for a skeleton for the self-consistent perturbation theory, the calculation contains the graphs:

$$\begin{aligned} \Sigma &= \text{diagram 1} - \text{diagram 2} + \text{diagram 3} \\ &= \text{diagram 4} + \text{diagram 5} + \text{diagram 6} + \text{diagram 7} + \dots \end{aligned} \quad (4.26)$$

This becomes evident, if the single iteration steps are carried out by use of Eq. (4.26) and the diagrammatic Dyson equation (2.12).

By noting  $G_{ij}(t, t')$  according to Eqs. (2.2) and (2.3), the renormalized graphs in the  $\omega$ -space may be given straightforwardly by the use of the rules of Section 2 [12]. The expressions for the graphs are more simplified and better illustrated by starting out from a different consideration described in the following.

The characteristic equation

$$G^0{}^{-1}|\varphi_i\rangle = 0; \quad |\varphi_i\rangle = \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix} \quad (4.27)$$

is renormalized:

$$G^{-1}|\varphi_{kl}\rangle = 0; \quad |\varphi_{kl}\rangle = C_{kl} \Sigma S_{ikl} |\varphi_l\rangle. \quad (4.28)$$

$C_{kl}$  being an arbitrary constant for the moment. The new "one-particle" functions  $\{\varphi_{kl}\}$  are not orthogonal. Choosing

$$|C_{kl}|^2 = P_{kl} \quad (4.29)$$

and considering Eqs. (3.9) and (3.5), one obtains the completeness relation

$$\sum_{k,l} |\varphi_{kl}\rangle \langle \varphi_{kl}| = 1. \quad (4.30)$$

Hence,  $\{\varphi_{kl}\}$  is a set of functions well adapted for the description of the renormalizing process  $G^0 \rightarrow G$ . Starting out from the skeleton we can describe the renormalization by very simple rules:

Only internal indices are transformed and following transformation properties hold:

$$\begin{aligned} k \in \{\text{occ}\} &\rightarrow (k, l) = \begin{cases} k \in \{\text{occ}\}, l = 0, -1, -2, \dots \\ k \notin \{\text{occ}\}, l = -1, -2, \dots \end{cases} \\ k \notin \{\text{occ}\} &\rightarrow (k, l) = \begin{cases} k \in \{\text{occ}\}, l = 1, 2, 3, \dots \\ k \notin \{\text{occ}\}, l = 0, 1, \dots \end{cases} \\ \varepsilon_k, k \in \{\text{occ}\}, &\rightarrow -I_{kl} \\ \varepsilon_k, k \notin \{\text{occ}\}, &\rightarrow -A_{kl} \end{aligned} \quad (4.31)$$

$$\begin{aligned} V_{aikkf} &= \langle \varphi_a(1) \varphi_i(2) | V(1,2) | \varphi_k(1) \varphi_f(2) \rangle \rightarrow \langle \varphi_a(1) \varphi_{i'}(2) | V(1,2) | \varphi_{k'}(1) \varphi_{f'}(2) \rangle \\ &= (P_{a1} P_{i1'} P_{k1''} P_{f1'''})^{1/2} \Sigma S_{ia1} S_{ji1'} S_{uk1''} S_{qf1'''} V_{ijuk} \equiv \bar{V}_{ii'kk'ff'}. \end{aligned}$$

The self-energy part given in Eq. (4.26) can be directly expressed:

$$\begin{aligned} \Sigma_{ij} &= \sum_{\sigma} (\bar{V}_{ikjk} - \bar{V}_{ikkk}) \\ &+ \sum_{\sigma_1} \frac{(\bar{V}_{ij'kf} - \bar{V}_{ij'fk}) \cdot \bar{V}_{jj'kf}}{\omega - I_{j'l} + A_{f'l} + A_{k'l''}} + \sum_{\sigma_2} \frac{(\bar{V}_{ij'kf} - \bar{V}_{ij'fk}) \bar{V}_{jj'kf}}{\omega - A_{j'l} + I_{j'l} + I_{k'l''}}. \end{aligned} \quad (4.32)$$

The sets of indices  $\sigma, \sigma_1, \sigma_2$  are to be taken from Eq. (4.31). As the self-consistent calculation of Eq. (4.32) is numerically too lengthy, it cannot be carried out without additional simplifications. By use of a finite number of unoccupied orbitals the sets of indices go up with the number of iteration steps. If the one-pole approximation is well satisfied (Section 3), we may write

$$\begin{aligned} \Sigma_{ij}^{(2)} &\approx \sum_{\substack{k, f \in \{\text{occ}\} \\ j' \notin \{\text{occ}\}}} \frac{P_{j'} P_k P_f (V_{ij'kf} - V_{ij'fk}) V_{jj'kf}}{\omega - A_{j'} + I_k + I_f} \\ &+ \sum_{\substack{k, f \notin \{\text{occ}\} \\ j' \in \{\text{occ}\}}} \frac{P_{j'} P_k P_f (V_{ij'kf} - V_{ij'fk}) V_{jj'kf}}{\omega - I_{j'} + A_k + A_f}. \end{aligned} \quad (4.33)$$

$P$ , being the pole strength of a main pole. The introduction of the one-pole approximation for the graph of first order yields the simple relation

$$\text{Diagram 1} - \text{Diagram 2} \approx \sum_{k \in \{\text{occ}\}} V_{ik} [j_k] (P_k - 1) \quad (4.34)$$

The numerical evaluation of this relation leads to small differences between big numbers. In order to get rid of this inaccuracy all constant graphs of third

*order are to be calculated explicitly.* Nevertheless, an interesting conclusion can be deduced from Eq. (4.34): the contribution of these graphs in the one-pole approximation is negative for  $i=j$ .

For the example of nitrogen the self-consistent calculation with the self-energy part given in Eq. (4.33) only slightly improves the results. The final results of the present perturbation theory for the VIPs of  $N_2$  are reported in Table 2.

## 7. Summary

The method of the Green's functions provides a new approach to the calculation of ionization potentials and electron affinities. Moreover, it makes it possible to investigate one-particle properties of the system. An integral equation relating the Green's functions with the self-energy part gives us freedom to collect certain graphs to infinite order with the help of the perturbation expansion of the self-energy part to finite order. The problem to evaluate the Green's functions has hence been reduced to an investigation of the graphs of the self-energy part.

Expanding the self-energy part to the first order of perturbation theory yields Koopmans' theorem. Already in the second order of the perturbation expansion additional poles of the Green's functions are obtained, which cannot be explained by Koopmans' theorem. These ionization potentials correspond to ion states the expansion of which in terms of electron configurations contains mainly such configurations which differ by more than one orbital from the configuration describing the ground state of the initial molecule in the one-particle picture. VIPs of this type and VIPs in their energy region are not considered here, but have been investigated and calculated elsewhere [12]. Here, only these poles of the Green's functions are investigated which are situated far from the poles of the self-energy part.

In order to find these poles, first of all the second order of the self-energy part has been taken into account [10-12]. The numerical results, however, *prove that the second order is far from being able to reproduce the experimental results.* By going over to the third order the results were improved, but then it was evident that it is useless to evaluate only a finite number of orders.

Theoretical considerations and numerical calculations revealed that it is *necessary to introduce an effective interaction as well as renormalized particle- and hole-lines* in order to estimate the contributions of the graphs of higher orders.

Applications: As mentioned above, the theory has been applied to closed-shell molecules successfully [11, 12, 15]. For open-shell systems the theory may be applied as well, but the question, if the approximations, used here, are reasonable, has still to be investigated.

One of the main advantages of this perturbation treatment, compared with the calculation of the total energies of the corresponding states, is that only a small correction term has to be calculated and therefore less accurate wavefunctions are sufficient to obtain satisfactory results. It is already shown that in most cases reasonable values for Koopmans' defect can be calculated even with semiempirical CNDO-wavefunctions [23, 25].

## Appendix A

For some examples the antigraph relation  $A3 \approx -A5$  is not as well satisfied as the other relations in Eq. (4.11). In case the contribution of  $A3 + A5$  is considerable,  $\omega$ -independent graphs have to be taken into account for the calculation of

(A1)

The estimation of these graphs may be performed analogously to the evaluation of the renormalized graph of second order.

*Acknowledgement.* I would like to thank the Deutsche Forschungsgemeinschaft for financial support, the Leibniz-Rechenzentrum of the Bavarian Academy of Sciences for the necessary computing time and Prof. Dr. G. Hohlneicher for helpful discussions.

## References

1. Brueckner, K.A., Gammel, J.L.: Phys. Rev. **109**, 1023 (1958). --- Brueckner, K.A., Masterson, Jr., K.S.: Phys. Rev. **128**, 2267 (1962)
2. Giell-Mann, M., Brueckner, K.A.: Phys. Rev. **106**, 364 (1957)
3. e.g., Schultz, T.D.: Quantum field theory and the many-body problem. New York: Gordon and Breach 1964. --- Schrieffer, J.R.: Theory of superconductivity. New York: Benjamin 1964. --- Schrieffer, J.R.: In: Phonons and phonon interactions, Ed. Bak Thor. New York: Benjamin 1964
4. Kelly, H.P.: Phys. Rev. **136**, B896 (1964)
5. Doll, J.D., Reinhardt, W.P.: J. Chem. Phys. **57**, 1169 (1972). Reinhardt, W.P., Doll, J.D.: J. Chem. Phys. **50**, 2767 (1969)
6. Cahse, R.L., Kelly, H.P.: Phys. Rev. A **6**, 2150 (1972)
7. Chang, E.S., Pu, R.T., Das, T.P.: Phys. Rev. **174**, 1 (1968)
8. Reinhardt, W.P., Smith, J.B.: J. Chem. Phys. **58**, 2148 (1973)
9. Miller, J.H., Kelly, H.P.: Phys. Rev. A **4**, 480 (1971)
10. Cederbaum, L.S., Hohlneicher, G., Peyerimhoff, S.: Chem. Phys. Lett. **11**, 421 (1971)
11. Cederbaum, L.S., Hohlneicher, G., v. Niessen, W.: Chem. Phys. Lett. **18**, 503 (1973)
12. Cederbaum, L.S.: Thesis, T. U. München 1972
13. Koopmans, T.: Physica **1**, 104 (1933)
14. Ecker, F., Hohlneicher, G.: Theoret. chim. Acta (Berl.) **25**, 289 (1972)
15. Cederbaum, L.S., Hohlneicher, G., v. Niessen, W.: Mol. Phys., in press
16. Meyer, W.: Inter. J. Quantum Chem. S **5**, 341 (1971)
17. Abrikosov, A., Gor'kov, L., Dzyaloshinskii, J.: Quantum field theoretical methods in statistical physics. Oxford: Pergamon Press 1965
18. Thouless, D.J.: The quantum mechanics of many-body systems. New York: Academic Press 1961
19. Mattuck, R.D.: Feynman diagrams in the many-body problem. London: McGraw-Hill 1967
20. Migdal, A.B.: Theory of finite Fermi systems and applications to atomic nuclei. London: John Wiley & Sons 1967
21. Kelly, H.P.: Phys. Rev. **131**, 684 (1963)
22. Kelly, H.P.: Phys. Rev. **144**, 39 (1966)
23. Kellerer, B., Cederbaum, L.S., Hohlneicher, G.: to be published
24. Cederbaum, L.S.: Mol. Phys., in press
25. Kellerer, B.: Thesis, T. U. München 1973

26. Pines, D.: The many-body problem. New York: Benjamin 1961
27. Brueckner, K. A.: Advan. Chem. Phys. **14**, 215 (1969)
28. Turner, D., Baker, C., Baker, A. B., Brandle, C.: Molecular photoelectron spectroscopy. London: Wiley-Interscience 1970
29. Albat, R., Gruen, N.: J. of Phys. B **6**, 601 (1973)
30. Albat, R.: Z. Naturforsch. **27a**, 545 (1972)

Dr. L. S. Cederbaum  
Physik-Department der  
Technischen Universität München  
Theoretische Physik  
8046 Garching bei München  
Reaktorgelände  
Federal Republic of Germany

# Modes of Interconversion in the Cycloheptene Ring

Giorgio Favini and Alberto Nava

Istituto di Chimica Fisica dell'Università, Via Saldini 50, Milano, Italy

Received April 9, 1973

The possible modes of interconversion for cycloheptene have been studied by carrying out calculations in the CNDO/2 and MINDO/2 approximations.

The transition state of lower energy is found along an asymmetrical pathway from the chair form, which is the most stable conformation, to the twist-boat form; the calculated activation energy agrees with the experimental data obtained from DNMR measures. The possibility of a change in the inversion mechanism for benzocycloheptene is also discussed.

**Key words:** Cycloheptene conformation – Benzocycloheptene, inversion of

## 1. Introduction

In cycloheptene there are two families of conformations. The chair family contains the regular symmetrical chair (C) which is relatively rigid and cannot undergo pseudorotation [1]. The boat family contains the regular symmetrical boat (B), the twist-boat (TB) having a  $C_2$  axis, and the biplanar form (II) in which carbon atoms  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_6$ , and  $C_7$  are coplanar (Fig. 1); the forms (II) and (TB) are intermediate in the pseudorotation itinerary between the pair of equivalent boat conformations (B and B\*).

Several modes of interconversion can be considered with different transition states (Fig. 1).

Our earlier calculations by Hendrickson's treatment [2] found the chair form favoured over the boat by 1.19 kcal/mole and over the twist-boat by 1.71 kcal/mole [3], while previously Pauncz and Ginsburg [4] using only a potential function for nonbonded H...H interactions calculated the boat to be more stable than the chair by 0.67 kcal/mole, and Allinger *et al.* [5] found the boat preferred by 0.94 kcal/mole.

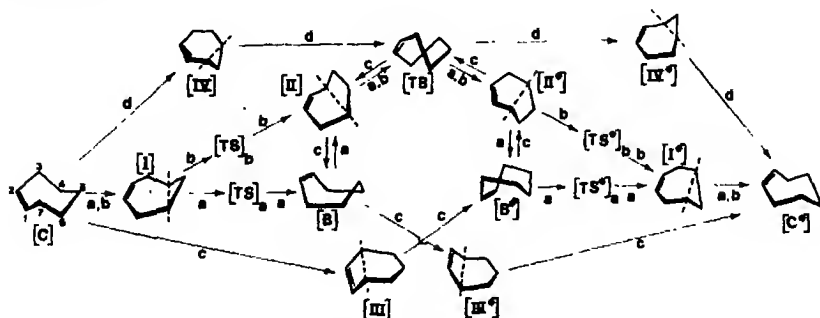


Fig. 1. Interconversion pathways between various forms of cycloheptene

Direct evidence for the structure of cycloheptene was obtained from a study of its Raman and infrared spectra [6]; it was concluded that the cycloheptene exists in the chair conformation in the crystal, while additional bands observed in the liquid at room temperature were attributed to the boat form. Dynamic nuclear magnetic resonance (DNMR) spectroscopy [7] has proven to be a powerful tool for the study of the ring inversion in cyclic molecules; St-Jacques and Vaziri [8] indicate that in 4,4,5,6,6-pentadeuteriocycloheptene there are two equivalent chair forms in equilibrium, separated by an energy barrier of  $5.0 \pm 0.3$  kcal/mole; the free energy profiles corresponding to the interconversion pathway are qualitatively discussed by the same authors.

In a more recent work Allinger and Sprague [9] with an extension of the force field method previously applied to saturated hydrocarbons, calculate the chair favoured by 0.57 kcal/mole over the twist form and by 3.37 kcal/mole over the boat form. Three interconversion pathways are also studied in detail; a symmetrical mode involves wagging of the  $C_7C_1C_2C_3$  plane through the form (I) having six coplanar carbon atoms (1, 2, 3, 4, 6, 7) and an energy barrier of 5.16 kcal/mole was calculated; a second symmetrical mode involves wagging of the  $C_4C_5C_6$  plane through the form (III) having five coplanar carbon atoms (3, 4, 5, 6, 7); the energy of this form was estimated to be of 10.26 kcal/mole; therefore it is an higher energy pathway contrary to suggestions of Kabuss *et al.* [10] and of Glazer *et al.* [11]. A third asymmetrical mode involves a direct interconversion between the chair and the twist-boat forms through an asymmetrical intermediate (IV) having six coplanar atoms (1, 2, 3, 4, 5, 7), for which an energy of 8.87 kcal/mole was evaluated [9].

Since the force field method requires a great amount of parameters and is founded on a classical basis, we will consider energetical and geometrical aspects of the possible modes of interconversion for cycloheptene within the framework of the molecular orbital method by carrying out all-valence electron SCF calculations in the CNDO/2 and MINDO/2 approximations.

## 2. Method of Calculation

The numbering system adopted is shown in Fig. 2. Bond lengths were assumed invariant as in the previous work [3]. Carbon atoms  $C_7$ ,  $C_1$ ,  $C_2$ , and  $C_3$  were considered always coplanar.

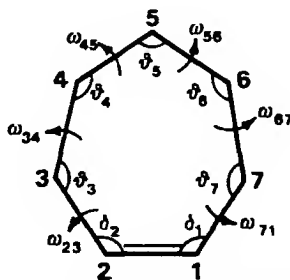


Fig. 2. Numbering system for cycloheptene



The following independently variable parameters were selected for the energy calculations:

- in the conformations with a plane of symmetry (C, B, I, III), the trigonal angle  $\delta_1 (= \delta_2)$ , the tetrahedral angle  $\vartheta_3$  and the dihedral angles  $\varphi$  and  $\varphi'$  (between the plane  $C_3C_4C_6C_7$  and the planes  $C_7C_1C_2C_3$  and  $C_4C_5C_6$  respectively).
- in the TB form, the angles  $\delta_1 (= \delta_2)$ ,  $\vartheta_4 (= \vartheta_6)$  and the dihedral angle between the planes  $C_7C_1C_2C_3C_5$  and  $C_3C_4C_5$  (or  $C_5C_6C_7$ ).
- in the biplanar form II, the angles  $\delta_1 (= \delta_2)$ ,  $\vartheta_3$  and  $\vartheta_5 (= \vartheta_6)$  and the dihedral angle between the two planes. With  $\vartheta_5 \neq \vartheta_6$  and  $\delta_1 \neq \delta_2$  the energy variation calculated are irrelevant.
- in the other asymmetrical forms, the angles  $\delta_1 (= \delta_2)$ ,  $\vartheta_3$ ,  $\vartheta_4$ ,  $\vartheta_5$ , and the torsional angles  $\omega_{23}$  and  $\omega_{34}$ .

The hydrogen atom positions were determined as functions of internal angles  $\vartheta$  by imposition of orthogonality between the hybrid orbitals on the carbon atom. The convention defined by Hendrickson [2] was used for the torsional angle signs.

The total energy was minimized with respect to couples of variables interpolating the function with an elliptical paraboloid. The CNDO/2 calculations were performed by employing the program QCPE 91 with usual parametrization; the MINDO calculations by modifying the program QCPE 137 (MINDO/1) in a MINDO/2 version with the parametrization indicated in Ref. [12].

### 3. Results

*Interconversion pathway a.* Symmetrical mode from the chair (C) to the boat (B) through the form (I) having six coplanar atoms and the transition state (TS)<sub>a</sub>, followed from pseudorotation to the twist-boat (TB) form. The potential function for this pathway is shown in schematic Fig. 3, and detailed in Fig. 4A for the rate

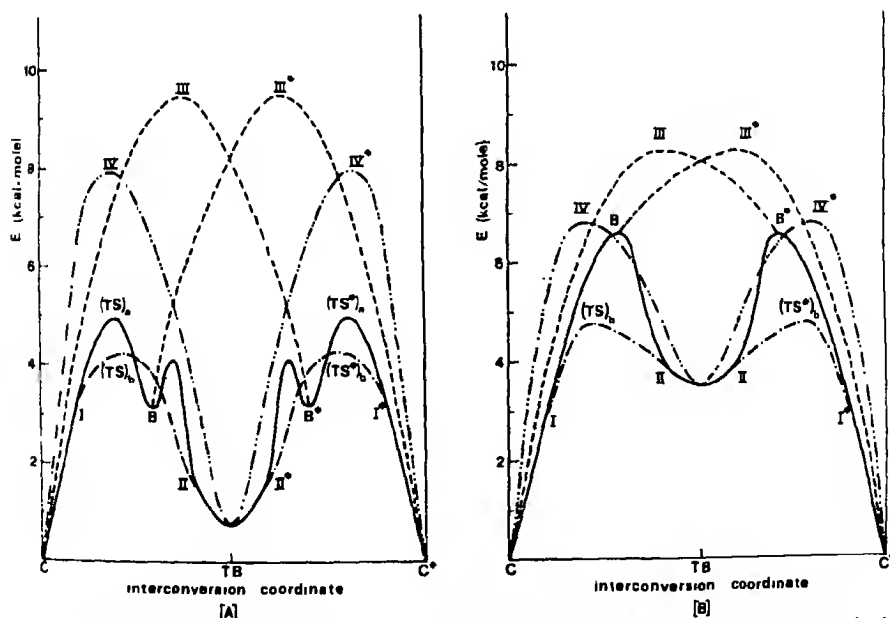


Fig. 3A and B. Schematic interconversion pathways. (A) CNDO/2 method. (B) MINDO/2 method.  
- pathway a, - - - pathway b, . . . . . pathway c, - · - · - pathway d

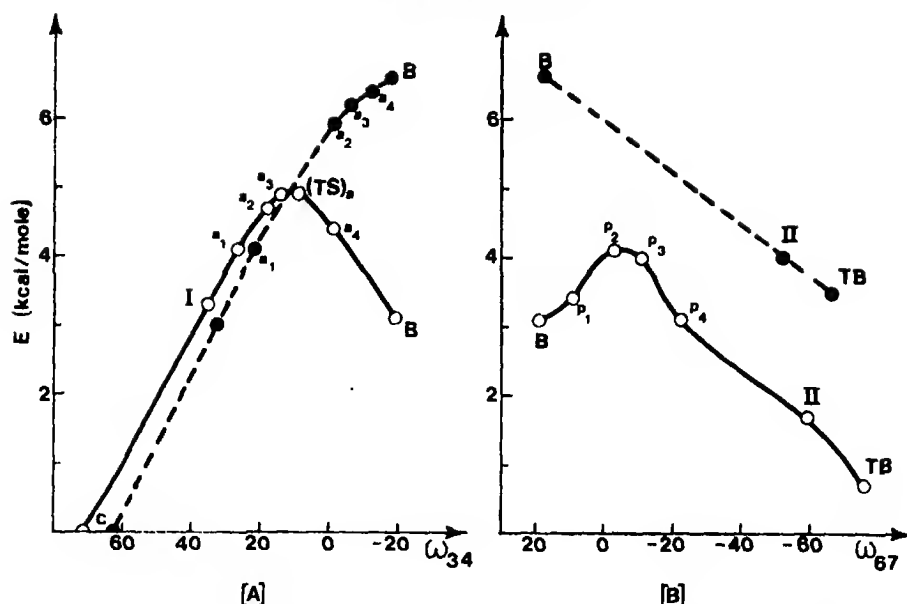


Fig. 4. (A), Symmetrical mode in pathway a; (B) Pseudorotation in pathway a. C'NDO 2 method, ● ● ● MINDO 2 method

Table 1 Geometry (angles in degrees) and energies (in kcal/mole) in the symmetrical mode of interconversion pathway a (C'NDO 2 method)

Form	(C)	(I)	$a_1$	$a_2$	$a_3$	(TS) <sub>a</sub>	$a_4$	(B)
$\theta_1 = \theta_2$	127.3	134.8	134.8	133.4	132.7	131.9	129.5	124.8
$\theta_3 = \theta_4$	113.0	118.5	118.0	118.0	117.5	117.0	116.0	113.0
$\theta_5 = \theta_6$	115.5	115.5	116.0	116.0	116.0	116.0	115.5	116.5
$\theta_5$	115.5	113.0	113.0	113.0	112.5	113.5	114.5	116.0
$\omega_{2,3} = \omega_{5,4}$	55.0	0.0	+ 11.0	+ 22.0	+ 27.5	+ 33.0	+ 44.0	+ 60.5
$\omega_{3,4} = \omega_{6,5}$	+ 71.5	+ 35.0	+ 26.5	+ 18.0	+ 14.0	+ 9.0	- 1.5	- 19.0
$\omega_{4,5} = \omega_{5,6}$	- 68.5	- 82.0	- 81.5	- 79.5	- 79.0	- 76.5	- 72.5	- 60.0
E	0.0	3.3	4.1	4.7	4.9	4.9	4.4	3.1

determining step (symmetrical mode) as function of the torsional angle  $\omega_{34}$  and in Fig. 4B for the pseudorotation step as function of the torsional angle  $\omega_{67}$ . Geometrical data and energies relative to the form (C), which is the most stable conformation, are reported in Tables 1-3. Angle values are rounded off to 0.5 degrees since smaller angular variations do not affect the energy significantly.

**Interconversion pathway b.** Symmetrical mode from the chair (C) to the form (I), followed by asymmetrical mode to the TB form through the transition state (TS)<sub>b</sub> (Fig. 3). The potential energies for this pathway are detailed in Fig. 5 as a function of  $\omega_{34}$  as the reaction coordinate. Geometrical data and energies of some conformations are reported in Table 4.

Table 2. Geometry and energies in the pseudorotation itinerary (CNDO/2 method)

Form	(B)	$p_1$	$p_2$	$p_3$	$p_4$	(II)	(TB)
$\delta_1 = \delta_2$	124.8	125.8	129.2	129.8	129.4	128.3	128.9
$\beta_3$	113.0	114.0	114.0	113.5	115.0	111.0	113.5
$\beta_4$	116.5	115.5	116.5	118.0	117.0	116.0	115.5
$\beta_5$	116.0	115.0	115.5	114.5	115.0	116.0	116.5
$\beta_6$	116.5	115.5	115.5	115.5	115.5	115.0	115.5
$\beta_7$	113.0	115.0	115.5	116.0	116.5	115.5	113.5
$\omega_{23}$	+ 60.5	+ 58.5	+ 51.0	+ 51.0	+ 53.5	+ 57.0	+ 32.0
$\omega_{34}$	- 19.0	- 17.5	- 12.0	- 15.5	25.0	- 66.5	- 76.0
$\omega_{45}$	- 60.0	- 63.0	- 65.5	61.0	52.0	0.0	+ 42.0
$\omega_{56}$	+ 60.0	+ 68.5	+ 75.0	+ 78.0	+ 82.0	+ 73.5	+ 42.0
$\omega_{67}$	+ 19.0	+ 9.0	- 3.0	- 11.0	- 22.0	- 59.5	- 76.0
$\omega_{-1}$	- 60.5	- 54.0	- 43.0	- 38.0	30.0	0.0	+ 32.0
$E$	3.1	3.4	4.1	4.0	3.1	1.7	0.7

Table 3. Geometry and energies in the interconversion pathway a (MINDO/2 method)

Form	(C)	(I)	$a_1$	$a_2$	$a_3$	$a_4$	(B)	(II)	(TB)
$\delta_1 = \delta_2$	127.4	132.3	132.3	129.5	128.5	127.0	125.6	127.9	128.2
$\beta_3$	120.5	122.5	122.5	121.5	121.0	120.5	120.0	118.0	120.0
$\beta_4$	118.5	119.0	119.5	120.0	120.0	119.5	119.5	118.5	119.0
$\beta_5$	116.0	113.5	115.5	117.0	119.0	119.0	119.5	118.5	119.0
$\beta_6$				= $\beta_4$				119.0	119.0
$\beta_7$				= $\beta_3$				120.5	120.0
$\omega_{23}$	- 44.0	0.0	+ 11.5	+ 35.0	+ 41.0	+ 47.5	+ 52.5	+ 51.0	+ 28.5
$\omega_{34}$	+ 62.5	+ 32.5	+ 22.0	- 1.0	- 6.0	12.0	- 18.0	- 60.5	- 66.0
$\omega_{45}$	- 62.0	- 72.0	- 69.0	- 61.5	- 56.5	- 53.5	- 49.0	0.0	+ 35.0
$\omega_{56}$				= - $\omega_{45}$				+ 62.5	+ 35.0
$\omega_{67}$				= - $\omega_{34}$				- 52.0	- 66.0
$\omega_{-1}$				= - $\omega_{23}$				0.0	+ 28.5
$E$	0.0	3.0	4.1	5.9	6.2	6.4	6.6	4.0	3.5

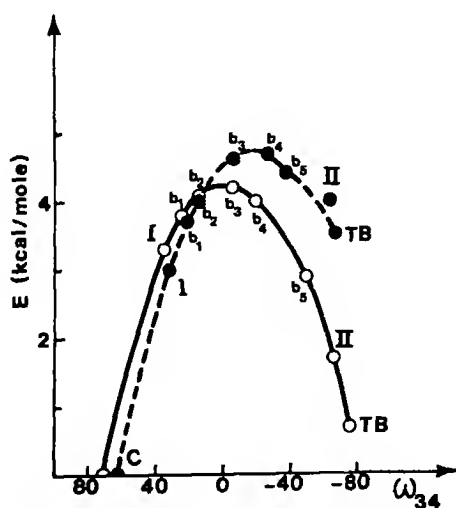


Fig. 5. Interconversion pathway b. - - - CNDO/2 method; ● - - - ● MINDO/2 method

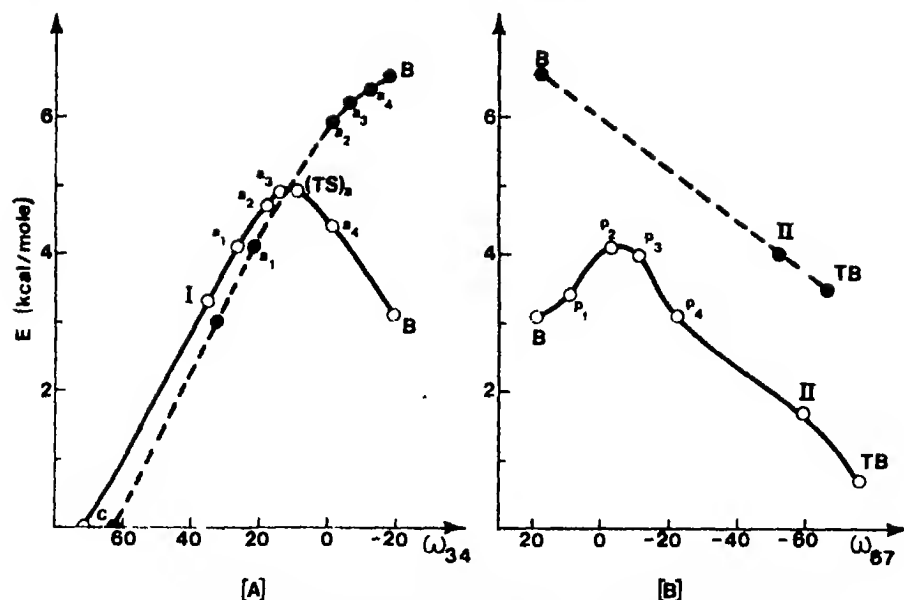


Fig. 4 (A), Symmetrical mode in pathway a; (B) Pseudorotation in pathway a.  
CINDO/2 method, ● ● ● MINDO/2 method

Table 1. Geometry (angles in degrees) and energies (in kcal/mole) in the symmetrical mode of interconversion pathway a (CINDO/2 method)

Form	(C)	(I)	$a_1$	$a_2$	$a_3$	(TS) <sub>a</sub>	$a_4$	(B)
$\theta_1$ $\theta_2$	127.3	134.8	134.8	133.4	132.7	131.9	129.5	124.8
$\theta_4$ $\theta_5$	113.0	118.5	118.0	118.0	117.5	117.0	116.0	113.0
$\theta_4$ $\theta_6$	115.5	115.5	116.0	116.0	116.0	116.0	115.5	116.5
$\theta_5$	115.5	113.0	113.0	113.0	112.5	113.5	114.5	116.0
$\omega_{13} = \omega_{45}$	55.0	0.0	+ 11.0	+ 22.0	+ 27.5	+ 33.0	+ 44.0	+ 60.5
$\omega_{14} = \omega_{65}$	+ 71.5	+ 35.0	+ 26.5	+ 18.0	+ 14.0	+ 9.0	1.5	- 19.0
$\omega_{45} = \omega_{46}$	68.5	82.0	- 81.5	- 79.5	79.0	- 76.5	- 72.5	- 60.0
E	0.0	3.3	4.1	4.7	4.9	4.9	4.4	3.1

determining step (symmetrical mode) as function of the torsional angle  $\omega_{34}$  and in Fig. 4B for the pseudorotation step as function of the torsional angle  $\omega_{67}$ . Geometrical data and energies relative to the form (C), which is the most stable conformation, are reported in Tables 1–3. Angle values are rounded off to 0.5 degrees since smaller angular variations do not affect the energy significantly.

**Interconversion pathway b.** Symmetrical mode from the chair (C) to the form (I), followed by asymmetrical mode to the TB form through the transition state (TS)<sub>b</sub> (Fig. 3). The potential energies for this pathway are detailed in Fig. 5 as a function of  $\omega_{34}$  as the reaction coordinate. Geometrical data and energies of some conformations are reported in Table 4.

Table 2. Geometry and energies in the pseudorotation itinerary (CNDO/2 method)

Form	(B)	$p_1$	$p_2$	$p_3$	$p_4$	(II)	(TB)
$\delta_1 = \delta_2$	124.8	125.8	129.2	129.8	129.4	128.3	128.9
$\vartheta_3$	113.0	114.0	114.0	113.5	115.0	111.0	113.5
$\vartheta_4$	116.5	115.5	116.5	118.0	117.0	116.0	115.5
$\vartheta_5$	116.0	115.0	115.5	114.5	115.0	116.0	116.5
$\vartheta_6$	116.5	115.5	115.5	115.5	115.5	115.0	115.5
$\vartheta_7$	113.0	115.0	115.5	116.0	116.5	115.5	113.5
$\omega_{23}$	+ 60.5	+ 58.5	+ 51.0	+ 51.0	+ 53.5	+ 57.0	+ 32.0
$\omega_{34}$	- 19.0	- 17.5	- 12.0	- 15.5	- 25.0	- 66.5	- 76.0
$\omega_{45}$	- 60.0	- 63.0	- 65.5	- 61.0	- 52.0	0.0	+ 42.0
$\omega_{56}$	+ 60.0	+ 68.5	+ 75.0	+ 78.0	+ 82.0	+ 73.5	+ 42.0
$\omega_{67}$	+ 19.0	+ 9.0	- 3.0	- 11.0	- 22.0	- 59.5	- 76.0
$\omega_{71}$	- 60.5	- 54.0	- 43.0	- 38.0	- 30.0	0.0	+ 32.0
$E$	3.1	3.4	4.1	4.0	3.1	1.7	0.7

Table 3. Geometry and energies in the interconversion pathway a (MINDO/2 method)

Form	(C)	(I)	$a_1$	$a_2$	$a_3$	$a_4$	(B)	(II)	(TB)
$\delta_1 = \delta_2$	127.4	132.3	132.3	129.5	128.5	127.0	125.6	127.9	128.2
$\vartheta_3$	120.5	122.5	122.5	121.5	121.0	120.5	120.0	118.0	120.0
$\vartheta_4$	118.5	119.0	119.5	120.0	120.0	119.5	119.5	118.5	119.0
$\vartheta_5$	116.0	113.5	115.5	117.0	119.0	119.0	119.5	118.5	119.0
$\vartheta_6$				$= \vartheta_4$				119.0	119.0
$\vartheta_7$				$= \vartheta_3$				120.5	120.0
$\omega_{23}$	- 44.0	0.0	+ 11.5	+ 35.0	+ 41.0	+ 47.5	+ 52.5	+ 51.0	+ 28.5
$\omega_{34}$	+ 62.5	+ 32.5	+ 22.0	- 1.0	- 6.0	- 12.0	- 18.0	- 60.5	- 66.0
$\omega_{45}$	- 62.0	- 72.0	- 69.0	- 61.5	- 56.5	- 53.5	- 49.0	0.0	+ 35.0
$\omega_{56}$				$= -\omega_{45}$				+ 62.5	+ 35.0
$\omega_{67}$				$= -\omega_{34}$				- 52.0	- 66.0
$\omega_{71}$				$= -\omega_{23}$				0.0	+ 28.5
$E$	0.0	3.0	4.1	5.9	6.2	6.4	6.6	4.0	3.5

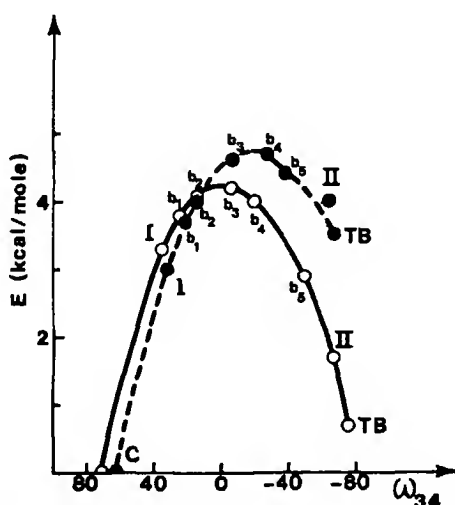


Fig. 5. Interconversion pathway b. --- CNDO/2 method; ●●● MINDO/2 method

Table 4. Geometry and energies in the interconversion pathway b

Form	CNDO/2 method				MINDO/2 method					
	$b_1$	$b_2$	$b_3$	$b_4$	$b_3$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$
$\delta_1 = \delta_2$	133.8	133.6	133.0	132.2	130.8	131.7	131.6	130.8	130.5	130.6
$\theta_1$	121.0	121.5	122.5	122.0	115.5	125.0	125.5	127.0	124.5	122.1
$\theta_4$	116.5	117.5	117.0	119.0	118.0	120.0	119.5	119.0	120.5	121.3
$\theta_5$	112.5	112.5	113.0	116.5	120.5	113.5	113.5	114.5	120.5	119.5
$\theta_6$	115.5	115.5	114.0	115.5	115.0	119.0	119.0	118.5	119.0	119.6
$\theta_7$	118.0	116.5	116.5	116.5	116.5	122.0	121.5	121.5	121.0	121.6
$\omega_{2,4}$	+ 7.5	+ 13.5	+ 20.0	+ 33.0	+ 48.0	+ 7.0	+ 11.0	+ 17.0	+ 33.5	+ 39.0
$\omega_{3,4}$	+ 24.0	+ 15.5	- 5.5	- 19.5	49.5	+ 22.0	+ 15.5	- 5.5	- 26.5	- 37.5
$\omega_{4,5}$	- 74.5	- 69.0	63.5	- 41.0	- 13.0	- 65.0	- 62.0	- 56.0	- 28.5	- 18.5
$\omega_{5,6}$	+ 85.5	+ 88.0	+ 90.5	+ 84.0	+ 74.0	+ 74.5	+ 77.0	+ 78.0	+ 69.0	+ 66.5
$\omega_{6,7}$	- 40.0	- 43.5	- 46.5	- 49.5	- 53.0	- 37.0	- 39.5	- 42.0	- 44.5	- 47.0
$\omega_{7,1}$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$E$	3.8	4.1	4.2	4.0	2.0	3.7	4.0	4.6	4.7	4.4

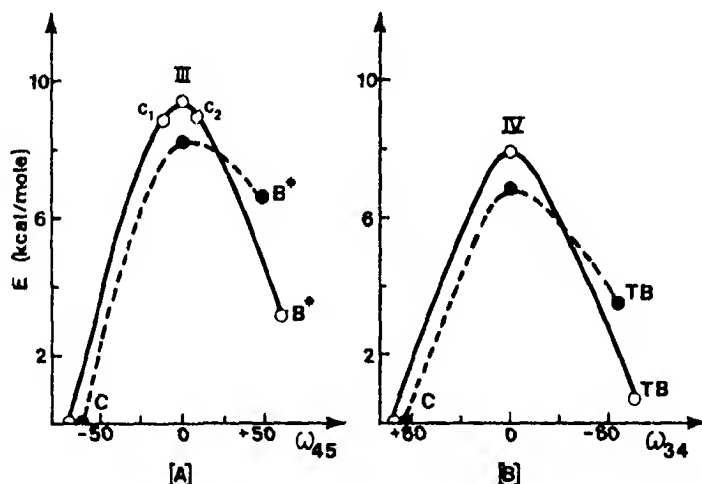


Fig. 6. (A) pathway c; (B) pathway d. — CNDO/2 method; --- MINDO/2 method

**Other Interconversion Pathways.** The symmetrical mode which involves wagging of the  $C_4C_5C_6$  plane through the form (III) having five coplanar carbon atoms (pathway c) and the asymmetrical mode with direct interconversion between the chair and the twist-boat forms through an intermediate (IV) having six coplanar atoms (atom 6 out of plane) (pathway d) are higher energy processes; the partial data obtained are shown in Fig. 6 and reported in Table 5.

#### 4. Discussion

The transition state of lower energy is found along the pathway b with both methods of calculation (Fig. 3). The calculated values of  $\Delta H^*$  for the chair inversion (4.25 kcal/mole by the CNDO/2 method and 4.75 kcal/mole by the MINDO/2 method) are slightly lower than the experimental value of  $\Delta G^* = 5.0$

Table 5. Geometry and energies in the interconversion pathways c and d

Form	CNDO/2 method				MINDO/2 method	
	$c_1$	(III)	$c_2$	(IV)	(III)	(IV)
$\delta_1 = \delta_2$	121.9	121.5	121.1	133.3	121.5	133.5
$\vartheta_3$	112.0	112.0	112.0	124.0	118.0	124.5
$\vartheta_4$	120.5	120.5	120.5	121.0	122.0	121.0
$\vartheta_5$	125.0	125.5	124.5	121.0	123.5	125.5
$\vartheta_6$	= $\vartheta_4$			112.5	= $\vartheta_4$	114.5
$\vartheta_7$	= $\vartheta_3$			110.5	= $\vartheta_3$	114.5
$\omega_{23}$	- 70.5	- 71.0	- 71.5	0.0	- 65.5	0.0
$\omega_{34}$	+ 58.5	+ 54.0	+ 49.0	0.0	+ 51.0	0.0
$\omega_{45}$	- 11.5	0.0	+ 11.5	- 35.5	0.0	- 33.0
$\omega_{56}$		= - $\omega_{45}$		+ 80.0	= - $\omega_{45}$	+ 71.0
$\omega_{67}$		= - $\omega_{34}$		- 72.0	= - $\omega_{34}$	- 63.5
$\omega_{71}$		= - $\omega_{23}$		+ 32.0	= - $\omega_{23}$	+ 29.0
E	8.8	9.4	8.9	7.9	8.2	6.8

$\pm 0.3$  kcal/mole found in pentadeuteriocycloheptene [8]. Since  $\Delta G^\ddagger$  and  $\Delta H^\ddagger$  are expected to be essentially equal, probably within experimental errors [8] (and indeed  $\Delta S^\ddagger = 0.2$  eu. in 5,5-difluorocycloheptene [11]), the agreement is very satisfactory. It is noteworthy that in proximity of the transition state (TS)<sub>b</sub> one observes a change in the sign of the dihedral angle  $\omega_{34}$  (Fig. 5); it means that there is a biplanar conformation with the atoms  $C_2C_3C_4C_5$  and  $C_6C_7C_1C_2C_3$  in two planes intersecting across the straight line  $C_2C_3$ .

The transition state (TS)<sub>a</sub> suggested by Allinger [9] is found only by the CNDO/2 method with energy somewhat higher (5 kcal/mole); its structure is about biplanar with atoms  $C_2C_3C_4C_5$  and  $C_5C_6C_7C_1$  in two planes with the atom  $C_5$  alone in common; also in the lower-energy pseudorotation process, the conformation with the highest energy value is characterized by a change in the sign of the torsional angle  $\omega_{67}$  (Fig. 4B). The MINDO/2 method, on the contrary, along the pathway a suggests that the boat form is the transition state, but it is reasonable to suppose that the inversion, after the form I is reached advances along the previously mentioned pathway b which is energetically preferred.

Also the biplanar intermediates (III) and (IV) are higher energy conformations and probably coincide with the transition states of the pathways c and d respectively.

As for the geometrical data, the small differences in the values calculated by the two methods are a consequence of the intrinsic properties of methods; really, higher bending constants are to be expected from CNDO/2 than from MINDO/2 and consequently the angles calculated for the minimum energy forms by the former method are generally smaller than those obtained from the latter method. A comparison with experimental data is not possible because they are not known, but the MINDO values should be the more reliable.

A recent proton magnetic resonance study of benzocycloheptene [12] showed unambiguously that the seven-membered ring exists as a chair conformation and a value of  $74\text{--}76^\circ$  was obtained for the torsional angle  $\omega_{34}$  from the coupling

constants. In the chair form of cycloheptene we calculate a value of  $71.5^\circ$  by the CNDO/2 method and of  $62.5^\circ$  by the MINDO/2 method; the agreement is fair when one takes into account that the trigonal angles in cycloheptene should be larger than in benzocycloheptene and consequently a diminution of torsional angles  $\omega_{34}$  and  $\omega_{67}$  is necessary for the ring closing.

The higher  $\Delta G^\ddagger$  value observed in the inversion of benzocycloheptene (10.9 kcal/mole [8]) was qualitatively interpreted previously [8, 9] on a different basis. We think that the replacement of the double bond by the benzene ring should stretch the bond  $C_1C_2$  and decrease the internal angles  $C_1\hat{C}_2C_3$  and  $C_2\hat{C}_1C_7$ . By going from the chair (C) to the transition states (TS)<sub>a</sub>, (TS)<sub>b</sub>, and (IV), a remarkable increase in the value of trigonal angles in the seven-membered ring occurs, whilst a decrease is observed in the transition state (III). Then in benzocycloheptene an increase in the inversion barrier along the pathways a, b, and d is very likely; the pathway c might become the most probable.

The possibility of this change in the inversion pathway would be able to account for the different values of  $\Delta G^\ddagger$  in some disubstituted benzocycloheptenes (10.3 kcal/mole in 3,3-dimethyl-, 11.3 kcal/mole in 4,4-dimethyl-, and 11.8 kcal/mole in 5,5-dimethyl-1,2-benzocycloheptene [10]). An increase of the internal tetrahedral angle on the carbon atom with geminal methyl groups brings an energetically unfavoured approaching of the substituents; in the transition state (TS)<sub>b</sub> we remark (Table 4) that  $\vartheta_3 > \vartheta_4 > \vartheta_5$ ; on the contrary, in the intermediate (III) the magnitude order (Table 5)  $\vartheta_5 > \vartheta_4 > \vartheta_3$  is in agreement with the observed free activation enthalpies.

We are planning to test if the same calculation methods can confirm this change of inversion mechanism in benzocycloheptene, in spite of the bigger molecular entity and of the higher energy values; preliminary results agree with our previous ones.

*Acknowledgement* Financial aid from Italian C. N. R. is gratefully acknowledged.

## References

1. Hendrickson, J. B.: *J. Am. Chem. Soc.* **89**, 7047 (1967)
2. Hendrickson, J. B.: *J. Am. Chem. Soc.* **83**, 4537 (1961)
3. Favini, G., Buemi, G., Raimondi, M.: *J. Mol. Struct.* **2**, 137 (1968)
4. Pauncz, R., Ginsburg, D.: *Tetrahedron* **9**, 40 (1960)
5. Allinger, N. L., Hirsch, J. A., Miller, M. A., Tyminski, I. J.: *J. Am. Chem. Soc.* **90**, 5773 (1968)
6. Neto, N., Di Lauro, C., Califano, S.: *Spectrochim. Acta A* **26**, 1489 (1970)
7. Binsch, G.: *Top. Stereochem.* **3**, 97 (1968)
8. St-Jacques, M., Vaziri, C.: *Canad. J. Chem.* **49**, 1256 (1971)
9. Allinger, N. L., Sprague, J. T.: *J. Am. Chem. Soc.* **94**, 5734 (1972)
10. Kabuss, S., Schmid, H. G., Friebohn, H., Faisst, W.: *Org. Magn. Resonance* **2**, 19 (1970)
11. Gilzer, E. S., Knorr, R., Ganter, C., Roberts, J. D.: *J. Am. Chem. Soc.* **94**, 6026 (1972)
12. St-Jacques, M., Vaziri, C.: *Org. Magn. Resonance* **4**, 77 (1972)

Prof. Giorgio Favini  
Istituto Chimica Fisica  
Università di Milano  
Via Saldini 50  
Milano, Italia



## On the Electronic Structure of $N_2H_2$ . A Possible Triplet Ground State in Diazines

G. Wagnière

Institute of Physical Chemistry of the University of Zurich\*

Received May 28, 1973

Within the frame of closed-shell and restricted open-shell *ab initio* SCF calculations 1,1-dihydrodiazine  $H_2N=N$  has a triplet ground state,  $^3A_2$ . This result, though not unsuspected from simple valence theory, is critically discussed and possible chemical implications are briefly mentioned.

**Key words:** Diazine Triplet ground state – Singlet-triplet separation – Cheletropic cleavage

### Introduction

The electronic structure of *cis*- and *trans*-diimide and of 1,1-dihydrodiazine has already been the subject of a number of theoretical investigations [1–6]. These molecules, though difficult or even impossible to isolate, are of particular interest for more than academic reasons: They are isoelectronic with such chemically important species as oxygen, formaldehyde and ethylene. *Cis*- and *trans*-diimide are the parent molecules of the chemically important class of the azo compounds [3, 7, 8]. The aminonitrenes, related to 1,1-dihydrodiazine, undergo cheletropic reactions liberating nitrogen [9–11]. Although all isomers of  $N_2H_2$  are thermodynamically unstable with respect to  $N_2$  and  $H_2$  [4], some spectroscopic data on  $N_2H_2$  is known [4, 12–14] which confirms the existence of the *trans* form.

In the present note we wish to report that within the frame of closed-shell and restricted open-shell SCF calculations [15] 1,1-dihydrodiazine has a triplet ground state. This coincides with the situation assumed to prevail in general in nitrenes [16]. We also compare the singlet-triplet splitting of the  $(n\pi^*)$  and  $(\pi\pi^*)$  excited states of the different  $N_2H_2$  isomers. The lowest excited states of *cis*- and *trans*-diimide are  $^3(n\pi^*)$ ,  $^1(n\pi^*)$ , and  $^3(\pi\pi^*)$ . In this context we will not consider the higher excited states, with the exception of  $^1(\pi\pi^*)$ .

### Calculations

The *ab initio* calculations were performed with the program IBMOL-5 [17] using a Gaussian-type basis. For all three isomers the geometry was considered as planar and, as in Ref. [3], the following values were adopted:  $r(NN) = 1.240 \text{ \AA}$ ,  $r(NH) = 1.021 \text{ \AA}$ ,  $\angle(NNH) = 112^\circ 39'$ . The orbital basis consisted for the calcula-

\* The *ab initio* calculations were performed at the IBM Research Center, San Jose, California.

tions listed under I in a nitrogen  $11s, 7p$  derived set and a hydrogen  $6s$  set [18] contracted to a  $(5, 3/2)$  set (see Tables 1a and 1b). The computations designated by II started from the same contracted basis, augmented by polarization functions in the plane of the molecule,  $p_x$  and  $p_y$ , of exponent  $\alpha = 1.000$  on each hydrogen atom, and by diffuse  $s, p_x, p_y, p_z$  functions of exponent  $\alpha = 0.030$  on each nitrogen atom.

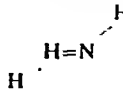
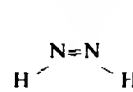
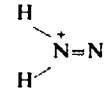
Table 1a. Orbital exponents of the uncontracted bases [18]

$s_N$	$p_N$	$s_H$
22800.2	82.7715	82.702833
3413.45	19.3820	12.418535
776.383	5.99713	2.825682
219.966	2.14355	0.798266
71.7952	0.820810	0.258165
25.8175	0.316980	0.089859
9.92954	0.118820	
3.94843		
1.11085		
0.436520		
0.160920		

Table 1b. Contraction coefficients listed in the sequence corresponding to Table 1a. Most values are taken from atomic SCF calculations [18]

	$s_N$		$p_N$		$s_H$
$\chi_1$	0.000221	$\chi'_1$	0.002359	$\chi''_1$	0.002004
	0.001715		0.017447		0.015333
	0.008908		0.074625		0.075518
	0.036023		0.212116		0.256733
	0.115506	$\chi'_2$	1.00000	$\chi'_2$	0.497645
	0.277557	$\chi'_3$	0.397308		0.296084
$\chi_2$	0.423141		0.167249		
	0.271130				
$\chi_3$	1.00000				
$\chi_4$	1.00000				
$\chi_5$	1.00000				

Table 2. Comparison of the total energy in a.u. of the lowest closed-shell SCF state

Calculation			
I	-109.96338	-109.94822	-109.94019
II	-109.98092	-109.96639	-109.95317
WFA GLO in situ* adjusted [4]	-109.93017	-109.92913	-109.92233
WFA Extended GLO* [4]	-109.95274	-109.94423	-109.94145

\* Geometry determined by energy minimization. See Ref. [4].

Table 3. Calculations I, II, and WFA [4] give energies in eV of the open-shell SCF states relative to the lowest closed-shell SCF state. The results listed under RHK [3] were obtained by a limited CI calculation

Calculation	State	$\begin{array}{c} \text{H} \\ \diagup \\ \text{N}=\text{N} \\ \diagdown \\ \text{H} \end{array}$		$\begin{array}{c} \text{H} \\ \diagdown \\ \text{N}=\text{N} \\ \diagup \\ \text{H} \end{array}$		$\begin{array}{c} \text{H} \\ \diagup \\ \text{N}^+=\text{N}^- \\ \diagdown \\ \text{H} \end{array}$	
I	$^3(n \rightarrow \pi^*)$	$^3B_g$	1.678	$^3B_2$	2.127	$^3A_2$	-0.171
	$^1(n \rightarrow \pi^*)$	$^1B_g$	3.017	$^1B_2$	3.659	$^1A_2$	1.273
	$^3(\pi \rightarrow \pi^*)$	$^3B_u$	3.697	$^3B_1$	3.656	$^3A_1$	5.566
	$^1(\pi \rightarrow \pi^*)$	$^1B_u$	12.435	$^1B_1$	12.441	$^1A_1$	5.732
II	$^3(n \rightarrow \pi^*)$		1.755		2.198		-0.090
	$^1(n \rightarrow \pi^*)$		3.077		3.722		1.328
	$^3(\pi \rightarrow \pi^*)$		3.689		3.651		5.430
	$^1(\pi \rightarrow \pi^*)$		11.263		11.372		5.593
RHK [3]	$^3(n \rightarrow \pi^*)$	$^3B_g$	3.01	$^3B_2$	2.07		
	$^1(n \rightarrow \pi^*)$	$^1B_g$	3.92	$^1B_2$	3.36		
	$^3(\pi \rightarrow \pi^*)$	$^3B_u$	6.53	$^3B_1$	6.45		
WFA [4]	$^3(n \rightarrow \pi^*)$	$^3B_g$	3.223				
	$^1(n \rightarrow \pi^*)$	$^1B_g$	3.838				

With the purpose of assessing the influence of methyl substitution on the energy of the  $(n\pi^*)$  states some semi-empirical SCF-CI calculations within the CNDO approximation are included (see Table 4). The difference between the parametrizations *a* and *b* concerns solely the core matrix elements. In case *a* they are determined following Wratten [19], in case *b* they are increased in their absolute value so as to depress the energy of the lower  $\sigma$  orbitals [8, 20].

### Results and Conclusions

In our calculation the geometry of the molecules was assumed fixed and no search for the minimum of the SCF energy with respect to the geometric parameters was undertaken. Nevertheless, with our orbital basis II the total ground state energy of *cis*- and *trans*-diimide turns out to be between 0.01 and 0.03 Hartrees lower than the better values listed in Ref. [4], which, though obtained using a Gaussian lobe basis, correspond to minima of the respective potential curves. We believe that minimization of the SCF energies with respect to geometry should not basically affect at least our qualitative conclusions. The finding that in  $\text{H}_2\text{N}^+=\text{N}^-$  the SCF energy of the  $^3A_2$  state lies below the one of the closed-shell  $^1A_1$  state is not unsuspected. This situation is in some sense analogous to the one encountered in the oxygen molecule. The  $2b_1(n)$  and  $2b_2(\pi^*)$  orbitals of the diazine are related to the highest filled  $e_{1g}(\pi^*)$  and  $\pi^*_z$  orbitals of oxygen (see Fig. 1). Because of the breakdown in cylindrical symmetry the degeneracy is of course lifted in the diazine. But the orbital energy gap is not as large as in the other  $\text{N}_2\text{H}_2$  isomers. The different environments of the two nitrogen atoms in the diazine leads to a relative localization of the molecular orbitals on one side or another of the molecule. In

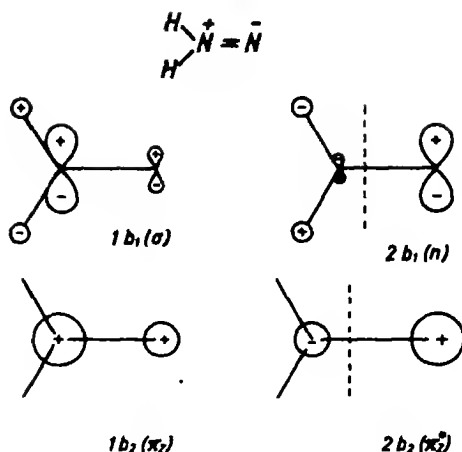


Fig. 1 Qualitative picture of orbitals for 1,1-dihydrodiazine

the lower  $1b_1$  orbital this localization takes place on the  $\begin{smallmatrix} \text{H} \\ \diagup \\ \text{N} \end{smallmatrix} > \text{N}$  side, as this favors H-N bonding. In  $2b_1$  the localization must, for orthogonality reasons, be on the other nitrogen. The very large orbital coefficients of  $2b_1(n)$  on this latter nitrogen atom, i.e., the "one-sidedness" of the MO, concomitantly reduces the antibonding effect of the node bisecting the N-N axis. The resulting molecular orbital may therefore be rated as quite characteristically nonbonding.

We notice that the improvement of the orbital basis, on going from I to II, reduces the SCF energy difference between  $^3A_2$  and  $^1A_1$ . Furthermore it cannot be ruled out that electron correlation will push  $^1A_1$  below  $^3A_2$ . Nevertheless it is conceivable that the  $^3A_2$  state might remain within thermal accessibility from the  $^1A_1$  ground state. However, we know of no simple way of assessing this possibly crucial difference in electron correlation energy.

The results of SCF-CI-CNDO calculations have been included for two reasons. Firstly, to assess the influence of methyl groups on the  $(n\pi^*)$  state energies. Secondly, to see in general where the average  $^3,1(n\pi^*)$  state energies come to lie with respect to the lowest closed-shell (ground) states. Within the ZDO approximation the singlet-triplet splitting cannot, of course, be computed. Table 4 shows these average energies to be very comparable to the *ab initio* SCF values, but in general to lie somewhat higher. As our semiempirical conclusions are incomplete without an attempt to estimate the magnitude of the singlet-triplet separation, and for the sake of a general comparison, we include the results of Table 5. The "simplified estimate" applies to  $(n\pi^*)$  and  $(\pi\pi^*)$  configurations described by MO's of the form:

$$\text{i) } \pi_z = \frac{1}{\sqrt{2(1+S)}} (p_{z1} + p_{z2}), \quad \pi_z^* = \frac{1}{\sqrt{2(1-S)}} (p_{z1} - p_{z2}),$$

$$n \equiv \pi_y^* = \frac{1}{\sqrt{2(1-S)}} (p_{y1} - p_{y2});$$

$$\text{ii) } \pi_z = p_{z1}, \quad \pi_z^* = p_{z2}, \quad n = p_{y2}.$$

Table 4. Energies in eV of the ( $n \rightarrow \pi^*$ ) states in  $N_2H_2$  and  $N_2(CH_3)_2$  as obtained by semiempirical SCF-CI-CNDO calculations. The singlet and triplet states are degenerate because of the ZDO approximation

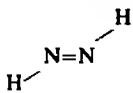
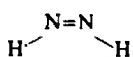

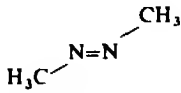
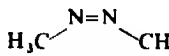
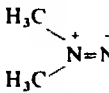
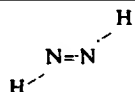
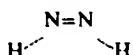

Parametrization			
<i>a</i>	3.10	4.07	1.77
<i>b</i>	2.90	3.83	1.43
			
<i>a</i>	2.37	3.21	1.94
<i>b</i>	2.29	3.14	1.59

Table 5. Singlet-triplet separation in eV. Values for the "simplified estimate" are based on functions as indicated in the text under i) and ii)

Calculation	State			
I	( $n \rightarrow \pi^*$ )	1.34	1.53	1.44
	( $\pi \rightarrow \pi^*$ )	8.74	8.78	0.17
II	( $n \rightarrow \pi^*$ )	1.32	1.52	1.42
	( $\pi \rightarrow \pi^*$ )	7.57	7.72	0.16
RHK [3]	( $n \rightarrow \pi^*$ )	0.91	1.29	
WFA [4]	( $n \rightarrow \pi^*$ )	0.62		
"simplified estimate" <sup>a</sup>	( $n \rightarrow \pi^*$ ) i)	1.98		ii) 2.24
	( $\pi \rightarrow \pi^*$ )	10.91		1.38

<sup>a</sup> Sahní, R. C., Cooley, J. W.: Tables for Molecular Integrals, TN D-146 II.

Use of the Mulliken approximation for integrals not listed therein.

The AO's are nitrogen Slater-type orbitals of exponent 1.950. If, as indicated, we take the results of Table 4 as averages of the respective singlet and triplet energies and apply our "simplified estimate" to then, then in 1,1-dihydrodiazine the lowest  $^3A_2$  state lies at about *a*) 0.65 eV, *b*) 0.31 eV above the closed-shell  $^1A_1$  state. For the dimethyl compound the corresponding values are *a*) 0.82 eV, *b*) 0.47 eV. This admittedly sketchy second procedure would indicate that the thermal accessibility of the lowest triplet state in diazines is rather unlikely.

The possible chemical consequence of a very low lying triplet state in diazines might be, for instance, that the cheletropic cleavage [9-11] of cyclic compounds

of the type shown below leads to diradical formation:



To determine unambiguously the course of such a reaction requires apparently much more elaborate computations on one hand, or suitable experiments on the other.

*Acknowledgement.* The author thanks the IBM Research Center, San Jose, for its hospitality and the IBM World Trade Corporation for financial support. He is grateful to Dr. Enrico Clementi for letting him use the program IBMOL-5 and for helpful discussions, and to Dr. Herbert Popkie for useful advice.

### References

1. Walsh, A. D.: *J. Chem. Soc.* **1953**, 2288
2. Wheland, G. W., Chen, P. S. K.: *J. Chem. Phys.* **24**, 67 (1956)
3. Robin, M. B., Hart, R. R., Kuebler, N. A.: *J. Am. Chem. Soc.* **89**, 1564 (1967)
4. Wong, D. P., Fink, W. H., Allen, L. C.: *J. Chem. Phys.* **52**, 6291 (1970)
5. Tinland, B.: *Spectrosc. Lett.* **3**, 51 (1970)
6. Bepalov, V. Y.: *Dokl. Akad. Nauk SSR* **200**, 99 (1971)
7. Haselbach, E., Heilbronner, E.: *Helv. Chim. Acta* **53**, 684 (1970)
8. Kuhn, J., Hug, W., Geiger, R., Wagnière, G.: *Helv. Chim. Acta* **54**, 2260 (1971)
9. Woodward, R. B., Hoffmann, R.: *Angew. Chem. Intern. Ed. Engl.* **8**, 781 (1969)
10. Lemal, D. M., McGregor, S. D.: *J. Am. Chem. Soc.* **88**, 1335 (1966)
11. Freeman, J. P., Graham, W. H.: *J. Am. Chem. Soc.* **89**, 1761 (1967)
12. Rosengren, K., Pimentel, G. C.: *J. Chem. Phys.* **43**, 507 (1965)
13. Trombetti, A.: *Canad. J. Phys.* **46**, 1005 (1968)
14. Boudybey, V. E., Nibler, J. W.: *Chem. Phys.* **58**, 2125 (1973)
15. Roothaan, C. C. J.: *Rev. Mod. Phys.* **32**, 179 (1960)
16. Gilchrist, T. L., Rees, C. W.: *Carbenes, nitrenes, and arynes*. London: Nelson 1969
17. Clementi, E., Mehl, J.: IBMOL-5 program user's guide, IBM research, RJ 889, 1971
18. Clementi, E., Mehl, J.: IBMOL-5 quantum chemical concepts and algorithms, IBM research, RJ 883, 1971, and private communication
19. Wratten, R. J.: *Chem. Phys. Letters* **1**, 667 (1968)
20. Hug, W., Kuhn, J., Seibold, K., Labhart, H., Wagnière, G.: *Helv. Chim. Acta* **54**, 1451 (1971)

Prof. Dr. G. Wagnière  
 Physikalisch-Chemisches Institut  
 der Universität Zürich  
 CH-8001 Zürich, Rämistrasse 76  
 Schweiz

## Calculations of Singlet and Triplet States of Some Azabenzenes by Modified INDO—CI\*

Sis-Yu Chen and Richard M. Hedges

Department of Chemistry, Texas A&M University  
College Station, Texas, USA 77843

Received November 13, 1972/May 29, 1973

Transition energies, ionization potentials, dipole moments and oscillator strengths have been calculated for pyridine, pyrazine, pyrimidine, pyridazine, 3-cyanopyridine and 4-cyanopyridine by modified INDO-CI. Triplet radiative lifetimes have been calculated for these molecules.

**Key words:** Ionization potential - Oscillator strength Spin-orbit coupling Azabenzenes

### 1. Introduction

Semi-empirical SCF all-valence-electron calculations to estimate transition energies of nitrogen-containing heterocyclic compounds have been performed by several authors. Del Bene and Jaffé's CNDO—CI calculations [1, 2] does not provide  $S-T$  splitting for  $n \rightarrow \pi^*$  transitions. Giessner-Prettre and Pullman [3] have used the INDO method with original parameterization [4] to calculate the transition energies of pyridine. The INDO method does give the  $S-T$  splitting, although the calculated values were too large. Yonezawa *et al.* [5] performed semi-empirical calculations on some azabenzenes taking all overlap integrals into account explicitly. *Ab initio* SCF MO calculations for the ground state of pyridine and pyrazine have been made by Clementi [6, 7] and by Petke *et al.* [8]. In the present work, using a modified INDO—CI method, the transition energies, oscillator strengths, ionization potentials and dipole moments are calculated and compared with existing experimental data.

Previous treatments of spin-orbit interaction in the singlet-triplet forbidden transitions of azines [9–12] always took the lone pair ( $n$ ) orbital as a localized hybrid orbital at the nitrogen atom. Yonezawa *et al.* [5] have studied the azabenzenes considering both of  $\sigma$  and  $\pi$  MO's explicitly. In the present work, using the spin-orbit coupling (SOC) perturbation on the INDO—CI wavefunctions, the singlet-triplet transition probabilities, the major sources of its intensity and the lifetimes of the triplet states are studied.

\* Taken from work presented to the Graduate College, Texas A&M University (by S.-Y. C.) in partial fulfillment of the requirements for the Ph. D. degree.

## 2. Calculations

### Ground State

The INDO method in its original parameterization was unable to give good transition energies without some refinement, and some ( $\sigma$ ,  $\pi$ ) transitions were inserted between the different ( $\pi$ ,  $\pi^*$ ) states. Our INDO calculations with some modifications of the original parameterization give improved results. These modifications are described as follows:

The monocenter coulombic integrals are the same those used by Yamaguchi and Fueno [13]. The parameters ( $\gamma_{AA}$ ) for hydrogen, carbon and nitrogen are 12.845, 11.715 and 12.860 respectively.

To evaluate the two-center coulombic integrals,  $\gamma_{AB}$ , we use a modified Mataga and Nishimoto type algorithm. The formula is

$$\gamma_{AB} = \langle \mu\mu | vv \rangle = \frac{e^2}{R_{AB}/0.529167 + a \cdot \exp(-bR_{AB}^c)}$$

where  $e$  is the electronic charge;  $R_{AB}$  is in Å units; the factor  $\exp(-bR_{AB}^c)$  is added to cause the integral curve to fall off with increasing distance  $R_{AB}$ ;  $b$  and  $c$  are determined for different atomic centers of A and B and are listed in Table 1. Parameter  $a$  is determined using the valence state ionization potential  $I_\mu$  and electron affinity  $A_\mu$  in the same valence state as follows:

(i) For the case of homonuclear interaction, we use  $a$  as did Mataga and Nishimoto.

(ii) For the case of heteronuclear interaction, we take the weighted average value,

$$a = \frac{\delta_A \gamma_{AA} + \delta_B \gamma_{BB}}{\delta_A + \delta_B}$$

where  $\delta$ 's are the atomic orbital exponents by Slater's rule.

The resonance integrals are taken the same as formulated by Jaffe [1].

$$\frac{1}{2} k (\beta_A^0 + \beta_B^0) S_{\mu\nu}$$

where the factor  $k$  is used to separate the  $\sigma$  and  $\pi$  type overlap integrals. For  $\sigma$  levels the value of  $k$  is unity, for  $\pi$  levels the value of  $k$  found more suitable in our calculation is 0.595. The bonding parameters  $\beta_A^0$  are 9 eV for hydrogen (the original CNDO value), 17 eV and 26 eV for carbon and nitrogen (the Jaffe's values) respectively.

The terms,  $-\frac{1}{2}(I_\mu + A_\mu)$ , of the core integrals are taken from Hinze and Jaffe's valence state data, except that the  $2p_z$  integral was decreased to decrease the stability of  $\pi$ -bonding; see Table 2.

Table 1. The constants in the exponential factor  $\exp(-bR^c)$  for computation of  $\gamma_{AB}$

Atomic center constants	N - N	N - C	N - H	C - C	C - H	H - H
$b$	0.8250	0.7964	0.8125	0.7713	0.7843	0.9730
$c$	0.7	0.68	0.7	0.7	0.7	0.5



Table 2. Some parameters in the INDO matrix elements (eV)

Atom	$-1/2 (I_p + A_p)$			$G^1$	$F^2$
	$s$	$p_{x,y}$	$p_z$		
C Original INDO	14.051	5.572	5.572	7.284	4.727
This work	14.960	5.805	5.000	6.180	3.886
N Original INDO	19.316	7.275	7.275	9.416	5.961
This work	19.536	7.550	6.250	7.750	5.070

Early attempts to utilize INDO with Jaffé's parameterization and the values of  $G^1$  and  $F^2$  given in original INDO paper [4] placed the  $n$  orbital higher than occupied  $\pi$  orbitals in pyridine molecule. This disagrees with the non-empirical results [6, 8] and the measured ionization potential definitely favours the  $\pi$  nature of the first ionization [19]. To overcome this difficulty, we used smaller values of  $G^1$  and  $F^2$  to increase the delocalization of  $n$  electrons and to get better results in spectroscopic properties. The values of  $G^1$  and  $F^2$  are given in Table 2.

### *Excited State Transitions and Intensities*

The excited states are generated by CI [14] among configurations formed by single electron excitation from occupied MO's into unoccupied or virtual orbitals determined in the ground state calculation. The number of configurations to be treated in the CI matrix is limited by a preset energy criterion. The energy criterion is usually between 9 and 10 eV or high enough to contain some ( $\sigma$ ,  $\pi$ ) transitions. Good agreement of calculated excited state energies with experiment is the criterion used to optimize the parametrization in the ground state basis functions.

The effect of SOC in organic systems is small enough to be treated by the first-order perturbation theory. McClure's [15, 16] central-field approximation is used in which the spin-orbit interaction for light atoms is expressed as:

$$H_{so} = \sum_i A_i l_i \cdot s_i$$

In the present treatment we shall need the expectation values of the spin-orbit interaction over the radial part of the  $2p$  atomic orbitals. The integral is:

$$\xi_k = \hbar^2 \langle A_{lk} \rangle_{av}.$$

The values of  $\xi_k$  for carbon and nitrogen are taken from Hartree-Fock atomic parameters [17] and are  $\xi_C = 31.95 \text{ cm}^{-1}$  and  $\xi_N = 75.59 \text{ cm}^{-1}$  respectively.

## **3. Results and Discussion**

### *Orbital Energies and Ionization Potentials*

The molecular orbital energies, bond orders and charge densities calculated by the present method are listed and discussed in the Chapter IV of Ref. [18]. The predicted molecular vertical ionization potentials (IP) via Koopmans'

Table 3. Calculated and observed dipole moments (Debyes)

Molecule	Calc.			Obs.
	$D_{ch}$	$D_A$	$D_I$	
Pyridine	0.91	1.73	2.64	2.21 <sup>a</sup>
Pyrazine	0.0	0.0	0.0	0.0
Pyrimidine	0.96	1.79	2.75	2.40 <sup>b</sup>
Pyridazine	1.83	3.27	5.10	3.97 <sup>b</sup>
3-cyanopyridine	2.33 <sub>x</sub>	2.06 <sub>x</sub>	4.39 <sub>x</sub>	-
	-0.37 <sub>y</sub>	0.83 <sub>y</sub>	0.46 <sub>y</sub>	
4-cyanopyridine	1.58	0.28	1.86	-

<sup>a</sup> Cumper, C. W. N., Vogel, A. I., Walker, S.: *J. Chem. Soc.* 3621 (1956)

<sup>b</sup> Landolt-Börnstein: *Zahlenwerte und Funktionen*, Bd. 1.3. Berlin-Göttingen-Heidelberg: Springer 1951

theorem are in reasonable agreement with the experimental data. The IP's of non-bonding electron for pyridine, pyrazine, pyrimidine and pyridazine are 10.89, 11.01, 10.76 and 11.00 eV comparing with the observed values 10.45, 10.15, 9.42 and 8.91 eV respectively. The IP's of non-bonding electron for 3-cyano and 4-cyano pyridines are 11.26 and 11.24 eV respectively.

### Dipole Moments

The total electronic dipole moments are divided into two parts,

$$D_I = D_{ch} + D_A$$

where the term  $D_{ch}$  represents the contribution due to the formal charges and the term  $D_A$  is the one form atomic polarization resulting from mixing of the 2s and 2p orbitals in same atom.

Our calculated values agree satisfactorily with the experimental data, and have the same relative ordering as obtains experimentally; see Table 3.

### Excitation Energies

Table 4 represents the calculated results for some lower excitation energies  $\Delta E$  and oscillator strengths  $f$ . The calculated lower  $n \rightarrow \pi^*$  and  $\pi \rightarrow \pi^*$  transition energies are in good agreement with experimental data except for the lowest  $n \rightarrow \pi^*$  transitions in pyrazine and pyridazine. The calculated singlet-triplet splitting in the lowest  $n \rightarrow \pi^*$  and  $\pi \rightarrow \pi^*$  transitions are reasonable in comparison with observed values.

Recently, Hoover and Kasha [20] concluded that the lowest triplet of pyridine is in fact  $n \rightarrow \pi^*$  and not  $\pi \rightarrow \pi^*$ , and from the phosphorescence in ethanol glass at 77 K, the lowest triplet was assigned to be  $^3A_1(n, \pi^*)$  for 3-cyanopyridine and  $^3B_1(n, \pi^*)$  for 4-cyanopyridine. This is in agreement with the results obtained in this work.

Table 4. Excitation energies  $\Delta E$  (eV) and oscillator strengths  $f$ 

Calc.			Obs.			Calc.			Obs.		
$\Delta E$		$f$	$\Delta E$		$f$	$\Delta E$		$f$	$\Delta E$		$f$
Pyridine						Pyrazine					
$^1B_1$	4.22	0.045	4.31	0.003 <sup>a</sup>		$^1B_{3u}$	3.28	0.08	3.83	0.01 <sup>c,a</sup>	
$^1B_2$	5.00	0.013	4.75	0.04 <sup>a</sup>		$^1B_{2u}$	4.61	0.05	4.69	0.1 <sup>a</sup>	
$^1A_1$	5.61	0.001	6.17	0.2 <sup>b,c</sup>		$^1B_{1u}$	5.39	0.0003	4.81	weak <sup>a</sup>	
$^1A_1$	7.26	0.01				$^1B_{1u}$	7.06	0.016	6.31	0.15 <sup>c,a</sup>	
$^1B_2$	7.68	0.55	6.82	1.3 <sup>b,c</sup>		$^1B_{2u}$	7.65	0.514	7.52	1.0 <sup>e</sup>	
$^1B_1$	7.69	0.083				$^3B_{3u}$	2.82		3.32 <sup>f</sup>		
$^3B_1$	3.58		3.68 <sup>d</sup>			$^3B_{2u}$	3.41				
$^3A_1$	3.66					$^3B_{1u}$	3.43				
$^3B_2$	4.12										
Pyrimidine						Pyridazine					
$^1B_1$	3.89	0.076	3.85	0.0069 <sup>c,e</sup>		$^1B_1$	3.92	0.115	3.30	0.0058 <sup>c,e</sup>	
$^1A_2$	5.19	0.0002				$^1A_1$	5.50	0.008	4.90	0.02 <sup>e</sup>	
$^1B_2$	5.21	0.01	5.00	0.052 <sup>e</sup>		$^1B_1$	5.65	0.019			
$^1A_1$	5.81	0.01	6.34	0.005 <sup>f</sup>		$^1B_2$	5.78	0.0038	6.2	0.10 <sup>e</sup>	
$^1B_1$	6.06	0.0031	6.49	0.016 <sup>e</sup>		$^1B_2$	7.52	0.22	7.10	1.0 <sup>f</sup>	
$^1A_1$	7.47	0.13	7.25	1.0 <sup>f</sup>		$^1A_1$	7.58	0.06			
$^1B_2$	7.50	0.0054				$^3B_1$	3.39		3.01 <sup>h</sup>		
$^1B_1$	3.56		3.63 <sup>a</sup>			$^3B_2$	3.81				
$^1A_2$	3.98					$^3A_2$	3.93				
$^1A_1$	4.02										
3-Cyanopyridine						4-Cyanopyridine					
$^1A''$	4.36	0.029	4.4 <sup>a</sup>	weak		$^1B_1$	4.03	0.028	3.99		
$^1A''$	4.93	0.0002				$^1B_2$	4.98	0.027	4.6 <sup>b</sup>	(2840)	
$^1A'$	5.01	0.019	4.7 <sup>i</sup>	(2230)		$^1A_1$	6.88	0.27			
$^1A'$	7.14	0.37				$^1B_2$	7.40	0.33			
$^1A'$	7.31	0.39				$^1B_1$	7.46	0.052			
$^3A'$	3.68		3.34 <sup>j</sup>			$^3B_1$	3.66	3.30 <sup>j</sup>			
$^1A''$	3.85					$^3A_1$	3.72				
$^3A'$	4.13					$^3B_2$	3.99				

Underline denotes  $\pi \rightarrow \pi^*$  type state; double underline denotes  $\sigma \rightarrow \pi^*$  type state; the rest are  $\pi \rightarrow \pi^*$  type states

<sup>a</sup>Sponer, H., Rush, J.: J. Chem. Phys. 20, 1847 (1952)

<sup>b</sup>Pickett, L. W., Corning, M. E., Wieder, G. M., Semenow, D. A., Buckley, J. M.: J. Am. Chem. Soc. 75, 1618 (1953)

<sup>c</sup>Kleven, H., Platt, J. R.: Univ. of Chicago Tech. Rept. (1953 - 54)

<sup>d</sup>Evans, D. F.: J. Chem. Soc. 3885 (1957)

<sup>e</sup>Mason, S. F.: J. Chem. Soc. 1240 (1959)

<sup>f</sup>Parkin, J. E., Innes, K. K.: J. Mol. Spectr. 15, 407 (1965)

<sup>g</sup>Goodman, L.: J. Chem. Phys. 46, 4731, 4737 (1967)

<sup>h</sup>Hochstrasser, R. M., Marzzacco, C.: J. Chem. Phys. 46, 1155 (1967)

<sup>i</sup>Jaffé, H. H., Orchin, M.: Theory and applications of ultraviolet spectroscopy. New York: John Wiley 1964

<sup>j</sup>See Ref. [20]

### Spin-orbit Coupling Mechanism

The contributions to the singlet-triplet transition moment by the perturbing singlet and triplet states are obtained by a spin-orbit coupling (SOC) perturbation calculation between singlet and triplet manifolds after CI and summed [18].

For the transition of  ${}^3B_1(n, \pi^*) \rightarrow S_0$  in the molecules having symmetry of the  $C_{2v}$  point group, such as pyridine, pyrimidine, pyridazine and 4-cyanopyridine, the perturbing singlet and triplet states are in the symmetry of  ${}^1B_2$ ,  ${}^1A_1$ ,  ${}^3A_2$  and  ${}^3B_1$ .

In  $D_{2h}$  symmetry (pyrazine), the perturbing states for  ${}^3B_{3u}(n, \pi^*) \rightarrow S_0$  are  ${}^1B_{1u}$ ,  ${}^1B_{3u}$ ,  ${}^3B_{3g}$  and  ${}^3B_{1g}$ . For the  ${}^3A'(\pi, \pi^*) \rightarrow S_0$  in 3-cyanopyridine, which has the symmetry of  $C_s$  point group, the perturbing states are  ${}^1A''$  and  ${}^3A''$ .

As is to be expected, the polarization of the transition  ${}^3(\pi, \pi^*) \rightarrow S_0$  is perpendicular to the molecular plane, while the polarization of  ${}^3(n, \pi^*) \rightarrow S_0$  is parallel to the molecular plane. For pyrazine and pyrimidine the transitions are polarized mainly along the N-N axes in agreement with experimental results [21, 22].

It is interesting to note that the contribution to the singlet-triplet transition is found to be from the component,  $r = +1$  and  $r = -1$ , of the triplet state wavefunction. No contribution to  $T_1 \rightarrow S_0$  is found from  $r = 0$  component of triplet states.

### The Radiative Lifetimes of Triplet State

The radiative lifetimes of triplet states are calculated by using the Mulliken approximation [23],

$$\tau = \frac{4.5}{\nu^2 \sum_r f^r} \quad (r = +1, -1 \text{ or } 0)$$

and are given in Table 5 along with experimental estimated values.

It is generally known that the values calculated from this approximation are not directly comparable to the observed values. Competing radiationless processes usually will cause the observed lifetime to be shorter than the calculated lifetime.

Table 5. Radiative lifetimes of the triplet states

Molecule	Triplet states	Calc. (sec.)	Obs. (sec.)
Pyridine	${}^3B_1(n, \pi^*)$	0.054 y	
Pyrazine	${}^3B_{3u}(n, \pi^*)$	0.052 y	0.02 <sup>a</sup>
Pyrimidine	${}^3B_1(n, \pi^*)$	0.072 y	0.01 – 0.02 <sup>b</sup>
Pyridazine	${}^3B_1(n, \pi^*)$	0.050 y	0.05 <sup>c</sup>
3-Cyanopyridine	${}^3A'(\pi, \pi^*)$	53.5 z	3.3 <sup>d</sup>
4-Cyanopyridine	${}^3B_1(n, \pi^*)$	0.038 y	0.004 <sup>d</sup>

<sup>a</sup>Goodman, L., Kasha, M.: J. Mol. Spectr. 2, 58 (1958)

<sup>b</sup>Goodman, L., Krishna, V. G.: Rev. Mod. Phys. 35, 541 (1963)

<sup>c</sup>Hochstrasser, R. M., Marzocco, C.: J. Chem. Phys. 46, 4155 (1967)

<sup>d</sup>See Ref. [20]

*Acknowledgements.* This work was supported by the Robert A. Welch Foundation. The calculations were carried out on the IBM 360/65 computer at The Data Processing Center of Texas A&M University.

### References

1. Del Bene, J., Jaffé, H. H.: *J. Chem. Phys.* **48**, 1807 (1968)
2. Del Bene, J.: *J. Chem. Phys.* **50**, 1126 (1969)
3. Giessner-Prettre, C., Pullman, A.: *Theoret. chim. Acta (Berl.)* **13**, 265 (1969)
4. Pople, J. A., Beveridge, D. L., Dobosh, P. A.: *J. Chem. Phys.* **47**, 2026 (1967)
5. Yonezawa, T., Kato, H.: *Theoret. chim. Acta (Berl.)* **13**, 123 (1969)
6. Clementi, E. J.: *J. Chem. Phys.* **46**, 4731 (1967)
7. Clementi, E. J.: *J. Chem. Phys.* **46**, 4737 (1967)
8. Petke, J. D., Whitten, L., Ryan, J. A.: *J. Chem. Phys.* **48**, 953 (1968)
9. El-Sayed, M. A.: *J. Chem. Phys.* **36**, 573 (1962)
10. Goodman, L., Krishna, V. G.: *Rev. Mod. Phys.* **35**, 54, 735 (1963)
11. Sidman, J. W.: *J. Mol. Spectr.* **2**, 333 (1958)
12. Vanquickenborne, L., McGlynn, S. P.: *J. Chem. Phys.* **45**, 4755 (1966)
13. Yamaguchi, K., Fueno, T.: *Bull. Chem. Soc. Japan* **44**, 43 (1971)
14. Pariser, R., Parr, R. G.: *J. Chem. Phys.* **21**, 767 (1953)
15. McClure, D. S.: *J. Chem. Phys.* **22**, 1668 (1954)
16. McClure, D. S.: *J. Chem. Phys.* **17**, 905 (1949)
17. Fischer, C. F.: Some Hartree-Fock results for the atoms helium to Radon. Univ. of British Columbia, Dept. of Math. 1968
18. Chen, S.-Y.: Dissertation, Texas A&M University, 1972
19. Goffart, C., Momigny, J., Natalis, P.: *Int. J. Mass Spect. Ion Phys.* **3**, 371 (1969)
20. Hoover, R. J., Kasha, M.: *J. Am. Chem. Soc.* **91**, 6508 (1969)
21. El-Sayed, M. A., Brewer, R. G.: *J. Chem. Phys.* **39**, 1623 (1963)
22. Krishna, V. G., Goodman, L.: *J. Chem. Phys.* **36**, 2217 (1962)
23. Mulliken, R. S.: *J. Chem. Phys.* **7**, 14 (1939)

Prof. Dr. R. M. Hedges  
Department of Chemistry  
Texas A&M University  
College of Science, College Station  
Texas 77843, USA



# An *ab initio* Molecular Orbital Study on the Addition Reaction of Triplet Nitrene to Ethylene

W. Joseph Haines and Imre G. Csizmadia

Lash Miller Chemical Laboratories, Department of Chemistry, University of Toronto,  
 Toronto M5S 1A1, Ontario

Received April 9, 1973

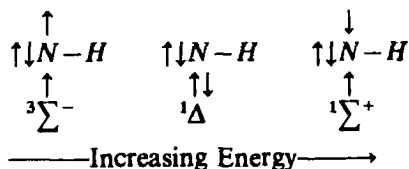
A minimal basis set, contracted from an extensive set of primitive Gaussian type functions (GTF), was used to expand the molecular orbitals (MO) within the framework of self consistent field (SCF) theory. The results revealed that aziridine is formed in its first excited triplet state ( $T_1$ ) when ethylene is reacted with triplet nitrene. The equilibrium geometry of aziridine in its  $T_1$  state had a tetrahedral CCN bond angle.

**Key words:** Aziridine    Addition reaction of triplet nitrene

## Introduction

Nitrenes have long been considered as important intermediates in organic chemistry. Much of the progress in nitrene chemistry has been made by investigation of the parallels between nitrene and carbene species. According to a qualitative theory, originated by Skell [1], the stereospecificity of carbene cycloaddition to olefins may be related to the multiplicity of the electronic state of the carbene molecule. The extension of Skell's method into nitrene chemistry has been previously studied experimentally by several authors [2, 3].

Any general nitrene species ( $R-N$ ) may exist in several electronic states which are iso-configurational in electronic structure. As an illustration one can consider the simplest nitrene ( $N-H$ ) whose three lowest electronic states are obtained by assigning the two electrons in the  $1\pi$ -shell to different spin orbitals.



Both experimentally and by SCF–MO calculations the triplet state is found to be the most stable species for  $N-H$  [4]. Electron spin resonance results also indicate that the triplet is the ground state for many other nitrenes [5].

The basic hypothesis advanced by Skell and adopted in the above experimental work [2, 3] is that (a) the singlet species should react in one single step process to yield the three-membered ring stereospecifically while (b) the triplet species should initially form a diradical intermediate which possesses

small barriers to rotations about single bonds with the result that by the time ring closure does occur all stereochemical integrity will have been lost. These ideas are well represented by Fig. 1.

The fundamental question we are primarily concerned in this communication is whether the first excited triplet state ( $T_1$ ) of aziridine is indeed in a ring distorted nuclear configuration as implied by Fig. 1.

### Computational Method

The Gaussian orbital basis used in this study was formed by variationally fitting sets of Gaussian type functions (GTF) to Slater type orbitals (STO) [6, 7]. As described and examined by Klessinger [8] these basis sets are not expected to give absolute energies approaching the Hartree-Fock limit but one hopes that these basis sets represent a near optimal compromise between computational efficiency and results which might parallel more exact energy surfaces.

In the basis set chosen each  $1s$  orbital was expanded in terms of 5 primitive Gaussians

$$\phi_{1s}(\xi, r) = \sum_{i=1}^5 c_i^{(1s)} \eta_{1s}(\xi^2 \alpha_i^{(1s)}, r).$$

Each  $2s$  Slater orbital and each  $2p$  Slater orbital was expanded in 3 primitive Gaussians.

$$\phi_{2s}(\xi, r) = \sum_{i=1}^3 c_i^{(2s)} \eta_{2s}(\xi^2 \alpha_i^{(2s)}, r)$$

$$\phi_{2p}(\xi, r) = \sum_{i=1}^3 c_i^{(2p)} \eta_{2p}(\xi^2 \alpha_i^{(2p)}, r)$$

The contraction coefficients and orbital exponents for the primitive Gaussians as determined for expansions of the above basis functions (BF) are those published by Klessinger [8]. For aziridine this represents 76 primitive GTF contracted to 20 BF.

The SCF-MO computations for species with closed electronic shells were performed according to Roothaan [9] and the energies of the singlet and triplet excited electronic configurations, involving the promotion of an electron from orbital  $b$  to orbital  $a$ , were calculated by the virtual orbital technique [9]

$${}^1E_{b \rightarrow a} = E_0 + (\epsilon_a - \epsilon_b) - J_{ab} + 2K_{ab}$$

$${}^3E_{b \rightarrow a} = E_0 + (\epsilon_a - \epsilon_b) - J_{ab}$$

where  $\epsilon_b$  and  $\epsilon_a$  are the occupied and virtual orbital energies respectively,  $J_{ab}$  is the coulomb and  $K_{ab}$  is the Exchange integrals over these ( $a$  and  $b$ ) molecular orbitals. The SCF MO calculations on nitrate were performed according to Roothaan's open shell formalism [10].

All SCF-MO computations were performed using the above basis set with the IBMOL-IV program system on an IBM 360/65 computer system at the University of Toronto.



### Results and Discussion

The calculations on the ground state of aziridine were performed at the experimental geometry determined by microwave spectroscopy [11], thus no attempt at optimization of geometry for this particular basis set was made. The value found for the total energy is somewhat higher than the results from the Double Zeta quality work [12, 13]. The character of the individual molecular orbitals is however the same as found in the more accurate calculations [12, 13].

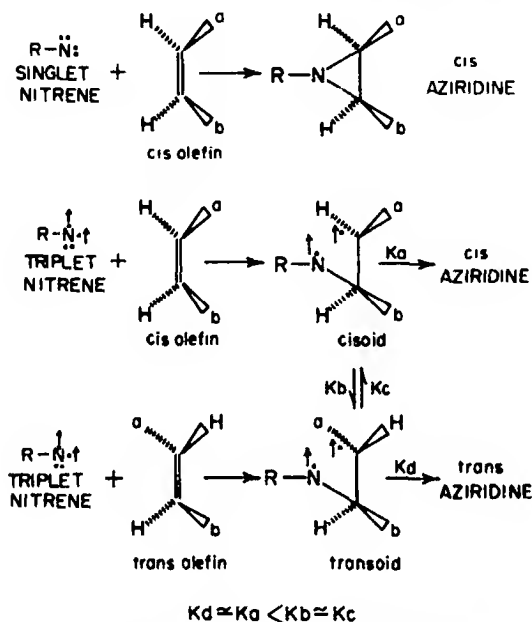


Fig. 1. Schematic description of nitrene addition to olefins

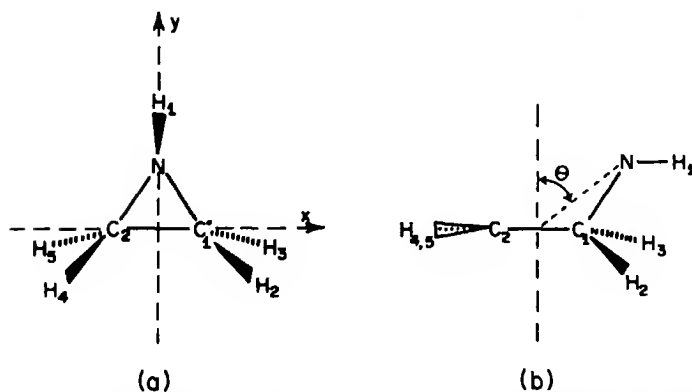


Fig. 2. (a) Coordinate system for aziridine in its equilibrium nuclear configuration  
(b) Illustration of the ring distortion in aziridine (The extent of ring distortion is measured by the angle  $\theta$ )

The highest occupied orbital, 8a, is predominately the nitrogen lone pair as evidenced from the contour plots in the plane containing the N-H bond and perpendicular to the CCN triangle.

The SCF-MO calculations on NH ( $^3\Sigma^-$  and  $^1\Delta$ ),  $C_2H_4$  ( $^1A_g$ ) and  $C_2H_5N$  ( $^1A_1$ ) yielded the following total energy values -54.75080, -54.62938, -77.74336 and -132.55322 hartree respectively. The first excited triplet state ( $T_1$ ) of aziridine was calculated by the virtual orbital technique [9] as -132.20843 hartree. These energy values allow one to draw a state, or more precisely a electronic configuration, correlation diagram (Fig. 3). As shown in Fig. 3 the cyclization is clearly favourable for the singlet NH reaction but there is a strong indication that the triplet reaction should procede through an intermediate of some distorted geometry which has a lower energy than the  $T_1$  in the ground equilibrium nuclear configuration. This ring distorted  $T_1$  species may then be deactivated to the  $S_0$  state without passing through such a highly unfavourable energy level as the one shown Fig. 3.

A partial geometrical optimization has been carried out on the ring distorted aziridine which is illustrated in Fig. 2b. Rotation about the CN bond showed that the proton attached to the nitrogen is now in the CCN plane and it points away from  $C_2$ . The CNH bond angle was found to be nearly tetrahedral similarly to the bond angles about  $C_1$ . The  $CH_2$  moiety involving  $C_2$  was chosen to have  $120^\circ$  bond angle (the HCH bond angles in aziridene are  $117^\circ$ ) and its plane was taken to be perpendicular to the CCN plane.

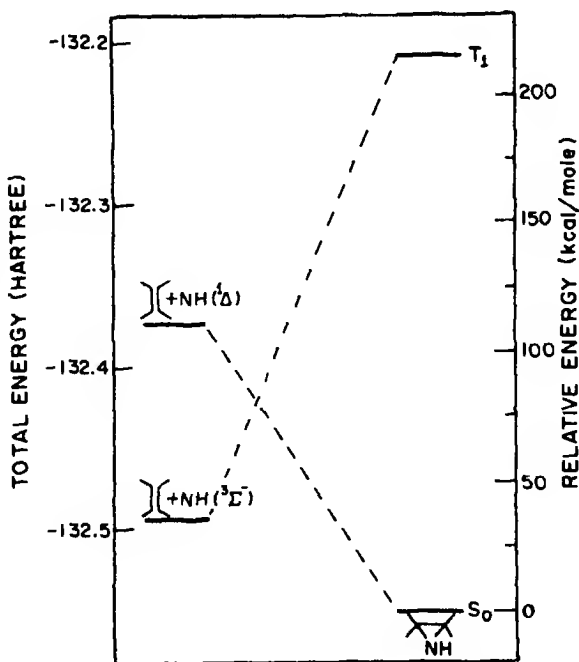


Fig. 3. Energy relationships between aziridine and its fragments

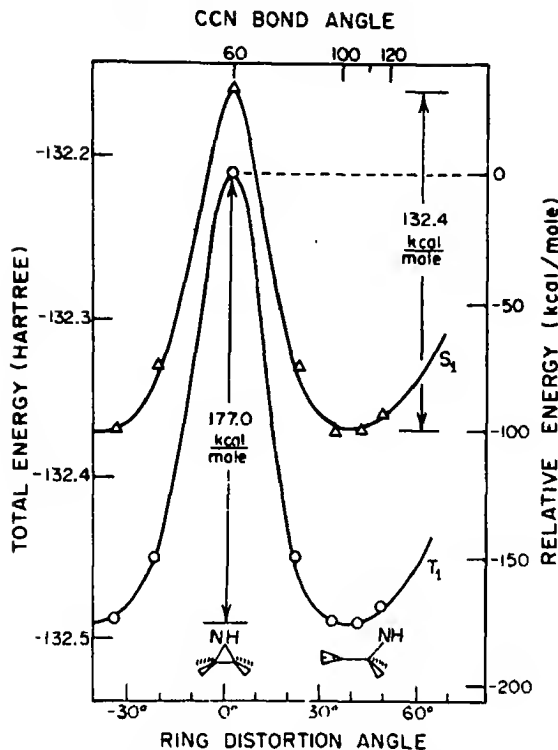


Fig. 4. Energy cross-section showing the effect of ring distortion upon the  $S_1$  and  $T_1$  states of aziridine

All further energy surface investigations were aimed at finding the minima for the first excited singlet and triplet ( $S_1$  and  $T_1$ ) states. The designation  $S_1$  and  $T_1$  will be most convenient because in many geometries considered no non-trivial elements of symmetry are present. Calculations were then performed at CCN angles of  $80^\circ$ ,  $100^\circ$ ,  $110^\circ$  and  $120^\circ$ . These points are shown in Fig. 4 where a ring distortion angle,  $\theta$ , (as illustrated in Fig. 2) is introduced for convenience. At each of these CCN angles the other bond angles were adjusted according to the previous paragraph. This result clearly indicates that in the first excited singlet and triplet states have the open (diradical) structure in contrast to the electronic ground state ( $S_0$ ) which has only one minimum at the equilibrium nuclear configuration (cf. Fig. 2).

This conclusion supports Skell's hypothesis that ethylene and triplet nitrene yields an open diradical structure in the primary step (cf. Fig. 1) and it is also in general agreement with the recent investigation of the ethylene and sulfur atom addition reaction [14].

Semi-quantitative studies by Hayes and Siu [15, 16] have considered symmetrical ring-open forms of aziridine ( $\text{CH}_2=\text{NH}-\text{CH}_2$ ) and postulated a 30% diradical character (as opposed to a resonating  $\pi$ -system  $\text{CH}_2=\dot{\text{N}}\text{H}-\bar{\text{C}}\text{H}_2\rightleftharpoons\bar{\text{C}}\text{H}_2-\dot{\text{N}}\text{H}=\text{CH}_2$ ) character for this open form. Salem

had previously demonstrated [17] that such diradical systems would have equal preference for concerted and non-concerted modes of ring closure and hence open forms with large diradical character should possess small degrees of stereochemical preference.

### References

1. Skell, P.S., Woodworth, R.C.: *J. Am. Chem. Soc.* **78**, 4496 (1956); Woodworth, R.C., Skell, P.S.: *J. Am. Chem. Soc.* **81**, 3383 (1959)
2. McConaghy, J.S.Jr., Lwowski, W.: *J. Am. Chem. Soc.* **88**, 2357 (1966)
3. Scheiner, P.: *J. Am. Chem. Soc.* **88**, 4759 (1966)
4. Huo, W.M.: *J. Chem. Phys.* **49**, 1482 (1968)
5. Wasserman, E., Smolinsky, G., Yager, W.A.: *J. Am. Chem. Soc.* **36**, 3166 (1964)
6. Huzinaga, S.: *J. Chem. Phys.* **42**, 1293 (1965)
7. Oohata, K., Taketa, H., Huzinaga, S.: *J. Physic. Soc. Japan* **21**, 2306 (1966)
8. Klessinger, M.: *Theoret. Chim. Acta (Berl.)* **15**, 353 (1969)
9. Roothaan, C.C.J.: *Rev. Mod. Phys.* **23**, 69 (1951)
10. Roothaan, C.C.J.: *Rev. Mod. Phys.* **32**, 1791 (1960)
11. Turner, T.F., Fiora, V.C., Kendrick, W.M.: *J. Chem. Phys.* **23**, 1966 (1955)
12. Lehn, J.M., Munsch, B., Millie, Ph., Veillard, A.: *Theoret. Chim. Acta (Berl.)* **13**, 313 (1969)
13. Basch, H., Robin, M.B., Kuebler, N.A., Baker, C., Turner, D.W.: *J. Chem. Phys.* **51**, 52 (1969)
14. Strausz, O.P., Gunning, H.E., Denes, A.S., Csizmadia, I.G.: *J. Am. Chem. Soc.* **94**, 8317 (1972)
15. Siu, A.K.O., St. John, 3rd, W.M., Hayes, E.F.: *J. Am. Chem. Soc.* **92**, 7249 (1970)
16. Hayes, E.F., Siu, A.K.Q.: *J. Am. Chem. Soc.* **93**, 2090 (1971)
17. Salem, L.: *J. Chem. Soc. Part D (Chem. Comm.)* **1970**, 981

Prof. Dr. J. G. Csizmadia  
Lash Miller Laboratories  
University of Toronto  
Dept. of Chemistry  
80 St. George St.  
Toronto 5, Canada

# On Bridging the Gap between the INDO and the NDDO Schemes

Björn Voigt

Department of Physical Chemistry, H. C. Ørsted Institute, University of Copenhagen,  
DK-2100 Copenhagen Ø, Denmark

Received March 20, 1973/June 4, 1973

A series of approximate LCAO SCF methods intermediate between the INDO and the NDDO schemes is proposed. The suggestion is based upon the decomposition of integrals in multipole-multipole type interactions.

**Key words:** Approximate LCAO SCF - Multipole-multipole interactions - INDO DRINDO - NDDO

## 1. Introduction

In recent years considerable efforts have been dedicated to studies of the CNDO (Complete Neglect of Differential Overlap) and the INDO (Intermediate NDO) approximate LCAO-SCF schemes [1–4]. On the other hand, only a few calculations have been performed using the less approximate NDDO (Neglect of Diatomic DO) scheme [1, 4]. Apparently the complications of the latter are judged to be disproportionately large compared with the gain expected in the reliability of the calculation. Most of the many additional integrals to be computed are unimportant, and they do not appear to be readily accessible to parameterization.

Of the few attempts made so far in bridging the gap between the INDO and the NDDO approximation schemes most have been plagued by the inadequacy of their not being invariant to the choice of local coordinate systems. The PNDO (Partial NDO) method [5] offers an example of this [4].

In the present paper a number of approximation schemes, all of which lie between the INDO and NDDO levels of approximation and all of which possess the desired invariance properties, is presented.

## 2. Proposals

The basis of the schemes to be proposed is the multipole expansion of the potentials from one-center charge distributions. To remove two-center charge distributions, the NDDO approximation is initially invoked. The remaining two-electron integrals are either integrals involving orbitals from only a single atom, or integrals which may be interpreted as representing the electrostatic interaction energy of two reasonably well separated one-center charge distributions ( $q_A$  and  $q_B$ ). The former integrals are all retained, as in the INDO method, while for the

latter the following expression is obtained when the usual expansion of  $1/r_{12}$  [6a] (around center A, say) is employed:

$$I(\varrho_A, \varrho_B) \equiv \int \int \varrho_A(r_1) \varrho_B(r_2) / r_{12} dv_1 dv_2 \quad (1)$$

$$= \sum_{l=0}^{\infty} \int \varrho_B(r_2) \left\{ \frac{4\pi}{2l+1} \int r_1^l \varrho_A(r_1) \sum_{m=-l}^{+l} Y_{lm}^*(\theta_1, \varphi_1) \frac{Y_{lm}(\theta_2, \varphi_2)}{r_1^{l+1}} dv_1 \right\} dv_2.$$

For non-overlapping charge distributions ( $r_1 = r_1$  and  $r_2 = r_2$ ), the expression in parentheses is the  $l^{\text{th}}$  term in the classical multipole expansion of the potential from  $\varrho_A$  [6b]. The fact that charge distributions do overlap in molecular integrals will not affect the argument given below.

When approximations are introduced into (1) the problem of coordinate invariance arises. The value of the integral should not depend upon the position of the origin and upon the orientation of the coordinate system chosen to describe the charge distributions and  $1/r_{12}$ . Since the integral is calculated exactly in the NDDO scheme, all coordinate systems will yield identical results in that case. For the INDO or the CNDO method, on the other hand, only the spherical term (the  $l = 0$ , or the monopole term) is retained. The integral is assigned a non-zero value only if the spherical average of  $\varrho_A$  (and of  $\varrho_B$ ) around the chosen origin differs from zero [7]. This truncation of the multipole expansion destroys the origin invariance. (This may easily be seen by considering the interaction of two point charges.) However, the problem is not serious when two atomic charge distributions are involved. In this case the natural choice for an origin is at either of the nuclear positions. It follows that computations with these two origins should yield the same result, and this implies in general, except for cases where the two charge distributions are identical and symmetrically arranged, that the integral must be calculated from the spherical averages of the two charge distributions around the respective centers. The orientation of the coordinate system employed in such a calculation is obviously immaterial; a spherical object looks the same in any direction from the center.

The approximations involved in the INDO method may be considered as an operator approximation for  $1/r_{12}$  suitably corrected to fulfill invariance requirements. However, one may adopt another point of view by expanding each of the charge distributions in spherical harmonics around the respective centers:

$$\varrho_A(r) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} R_{lm}(r) Y_{lm}(\theta, \varphi) \equiv \sum_{l=0}^{\infty} \varrho_A^{(l)}(r) \quad (2)$$

and similarly for  $\varrho_B$ . Inserting these expansions into (1) one obtains

$$I(\varrho_A, \varrho_B) = \sum_{k,l} \int \int \varrho_A^{(l)}(r_1) \varrho_B^{(k)}(r_2) / r_{12} dv_1 dv_2 \equiv \sum_{k,l} I_{kl} \quad (3)$$

with

$$I_{kl} = \int \varrho_B^{(k)}(r_2) \left\{ \frac{4\pi}{2l+1} \int r_1^l \varrho_A^{(l)}(r_1) \sum_{m=-l}^{+l} Y_{lm}^*(\theta_1, \varphi_1) \frac{Y_{lm}(\theta_2, \varphi_2)}{r_1^{l+1}} dv_1 \right\} dv_2. \quad (4)$$

The two-center interactions retained in the INDO scheme,  $I_{00}$ , may be characterized as monopole-monopole types. However, the expansion of integrals in

multipole-multipole type interactions may be continued to any order without loss of coordinate invariance. This follows directly from the addition theorem for spherical harmonics [6a]:

$$P_l(\cos \alpha) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_{lm}^*(\theta_1, \varphi_1) Y_{lm}(\theta_2, \varphi_2). \quad (5)$$

The angle  $\alpha$  between the radius vectors of the two points does not depend upon the coordinate system as long as the origin is fixed. Thus by inserting (5) into (4) all explicit reference to the coordinate system is removed. Each term in the double sum in (3) is therefore coordinate system invariant and the summation itself may be truncated in any manner. Of course, in molecular calculations the truncation must be symmetric in the sense that inclusion of  $I_{kl}$  implies that  $I_{lk}$  also be included. Thus the interactions retained are classified according to the multipole-multipole type.

The above analysis leads to the following proposal for the assignment of values to two-center integrals:

Firstly, the total charge of each of the charge distributions is calculated. The corresponding monopole-monopole type interaction energy, defined in such a way that it depends only upon the nature of the two atoms, is then computed. In addition, an analysis is made of whether or not the charge distributions have any multipole moments relative to their respective centers. For any that have the corresponding multipole-multipole (*i.e.* monopole-dipole, dipole-dipole, monopole-quadrupole, *etc.*) interaction energies of given types are added. These should, of course, depend upon the relative orientation of the multipoles, but otherwise only upon the atoms involved and the distance between them.

Such a procedure will be coordinate invariant. It may be shown to be invariant under hybridization transformations as well [3].

When the detailed form of the atomic charge distributions is taken into account it can be seen that for a finite basis set the expansion in (2) must terminate at some finite  $l$ . For a basis set containing *s*- and *p*-orbitals only, the highest non-vanishing term that may occur in (2) is the quadrupole term ( $l=2$ ). This is easily seen using the coupling rule for spherical harmonics [8]. Thus a calculation retaining interactions up to and including the quadrupole-quadrupole type in this case would be equivalent to the NDDO scheme. As a consequence, the decomposition of integrals into multipole-multipole interactions will give rise to only a finite number of approximation schemes that are intermediate between the INDO and the NDDO schemes.

### 3. The DRINDO Scheme

In the simplest of the approximate LCAO-SCF schemes proposed above monopole-dipole interactions are included in addition to the monopole-monopole terms. This scheme may be given the name DRINDO (Dipoles Retained INDO). The field which is made self-consistent in the DRINDO scheme is con-

structed taking into account the main effects of the polarization of atoms in molecular environments.

In the following a basis set of unhybridized atomic orbitals (AO's) consisting only of *s*- and *p*-orbitals is assumed. The symbols  $\nu$  and  $\lambda$  denote general AO's, while *s*,  $\sigma$  and *i* (or *j*) denote specific AO's. For atom A,  $s_A$  is the *s*-orbital,  $\sigma_{AB}$  is the *p*-orbital pointing towards atom B, and  $i_A$ ,  $i = x, y, z$ , are the *p*-orbitals directed along the axis of the local coordinate system on A.

With this basis set, the two-electron, two-center integrals which, according to the DRINDO procedure, should be retained are:

$$\begin{aligned} [\nu_A \nu_A | \lambda_B \lambda_B] &= \gamma_{AB} \\ [s_A i_A | \lambda_B \lambda_B] &= \delta_A^B (e_{Ai} \cdot e_{AB}) \end{aligned} \quad (6)$$

$e_{Ai}$  and  $e_{AB}$  are unit vectors directed along the *i*-axis on A and from A towards B respectively. If it is assumed that the *s*- and *p*-orbitals have the same radial dependence the expression for  $\delta_A^B$  is:

$$\delta_A^B = [s_A \sigma_{AB} | s_B s_B].$$

As noted by Dixon [9], it would be inconsistent to include the integrals (6) while neglecting the monopole-dipole interactions with the atomic cores. Thus the integrals

$$\langle s_A | V_B | i_A \rangle = d_A^B (e_{Ai} \cdot e_{AB})$$

where  $V_B$  is the attractive core potential from atom B, must also be included,  $d_A^B$  being given by

$$d_A^B = \langle s_A | V_B | \sigma_{AB} \rangle.$$

The matrix elements of the Fock operator in the DRINDO approximation are now readily constructed. Using superscripts to indicate the centers of the orbitals the following expressions are obtained for the one-center elements:

$$\begin{aligned} F_{\nu\nu}^{AA} &= F_{\nu\nu}^{\text{INDO}} + \sum_{B \neq A} 2\delta_A^B \left( e_{BA} \cdot \sum_j P_{sj}^{BB} e_{Bj} \right), \\ F_{si}^{AA} &= F_{si}^{\text{INDO}} + e_{Ai} \cdot \sum_{B \neq A} e_{AB} (P_{BB} \delta_A^B + d_A^B). \end{aligned}$$

Here  $F_{\nu\lambda}^{\text{INDO}}$  is the corresponding matrix element in the INDO approximation,  $P_{\nu\lambda}^{BB}$  denotes an element of the charge and bond order matrix, while the gross electronic population  $P_{BB}$  on atom B is defined as:

$$P_{BB} = \sum_{\lambda} P_{\lambda\lambda}^{BB}.$$

The two-center matrix elements are given in a similar notation by:

$$\begin{aligned} F_{ss}^{AB} &= F_{ss}^{\text{INDO}} - \frac{1}{2} \delta_A^B (e_{AB} \cdot \sum_i P_{is}^{AB} e_{Ai}) \\ &\quad - \frac{1}{2} \delta_B^A (e_{BA} \cdot \sum_j P_{sj}^{AB} e_{Bj}) \end{aligned}$$



$$\begin{aligned}
 F_{sj}^{AB} &= F_{sj}^{\text{INDO}} - \frac{1}{2} \delta_A^B (e_{AB} \cdot \sum_i P_{ij}^{AB} e_{Ai}) \\
 &\quad - \frac{1}{2} \delta_B^A P_{ss}^{AB} (e_{BA} \cdot e_{Bj}) \\
 F_{ij}^{AB} &= F_{ij}^{\text{INDO}} - \frac{1}{2} \delta_A^B P_{sj}^{AB} (e_{AB} \cdot e_{Ai}) \\
 &\quad - \frac{1}{2} \delta_B^A P_{is}^{AB} (e_{BA} \cdot e_{Bj}).
 \end{aligned}$$

The above expressions do not describe the manner in which the result of corresponding INDO and DRINDO calculations differ. However, an investigation of how two atoms, one of which is a hydrogen atom, interact in the two schemes does yield informations on this point. It is well-known that the total energy expression in the INDO approximation can be partitioned into monatomic and diatomic contributions [3]. This is retained in the DRINDO scheme. Since the additional integrals included in the DRINDO approximation all involve two centers, it follows that the monatomic terms are unchanged. If  $h$  denotes the hydrogen  $1s$ -orbital, the difference between the diatomic contributions from the atom pair A and H is:

$$E_{AH}^{\text{DRINDO}} - E_{AH}^{\text{INDO}} = 2P_{s\sigma}^{AA} \left\{ P_{HH}[hh|s\sigma] - \langle s | \frac{e^2}{r_H} | \sigma \rangle \right\} - P_{sh}^{AH} P_{\sigma h}^{AH} [hh|s\sigma]. \quad (7)$$

The charge distributions ( $hh$ ) and  $2P_{s\sigma}^{AA}$  can, for the present discussion, be regarded as a point charge and a point dipole respectively. Thus the first term on the right-hand side describes the electrostatic interaction between the net charge on the hydrogen atom and the effective dipole on A.

The last term in (7) is a bond term of second order. This term will be important only if the two atoms are bound together. In a localized description where both bond orders involved may be derived from a single molecular orbital describing the bond we have:

$$P_{sh}^{AH} P_{\sigma h}^{AH} = P_{HH} P_{s\sigma}^{\text{bond}}. \quad (8)$$

The bond term thus removes from the electrostatic term half of the interaction between the electronic charge on hydrogen and that part of the dipole on A which is due to the two bonding electrons. The electron does not interact with that portion of the dipole which it has produced itself.

For atoms far apart, the difference (7) is dominated by the pure electrostatic interaction. For atoms bound together, however, the situation is reversed. In this case the bond term which involves the gross electronic population rather than the net charge will usually play the major role. This term describes an attraction between the two atoms which, according to (8), increases with increasing  $sp$ -mixing in the bonding hybrid.

The above analysis indicates that the results of a DRINDO calculation will deviate from those from the corresponding INDO calculation by having:

1. Enhanced hybridization and thus increased polarization of atoms.
2. Smaller computed bond distances, in particular for short bonds (AH-bonds) where monopole-dipole interactions are largest.
3. Larger electron density on atoms bound to polarized (highly electronegative) atoms such as nitrogen, oxygen and fluorine. This will lower the electrostatic

repulsion between the atoms by roughly twice the amount with which the bond attraction is decreased.

The quantities  $\delta_A^B$  and  $d_A^B$  which constitute the difference between the DRINDO and the INDO schemes may be treated in various ways. They are easily calculated theoretically [10], but they may also be parameterized semiempirically. For example, corresponding to the Ohno-Klopman expression [11, 12]:

$$\gamma_{AB} = e^2 [R_{AB}^2 + (\varrho_A + \varrho_B)^2]^{-\frac{1}{2}}$$

one may use

$$\delta_A^B = e^2 \{ [(R_{AB} - A_A)^2 + (\varrho_A + \varrho_B)^2]^{-\frac{1}{2}} - [(R_{AB} + A_A)^2 + (\varrho_A + \varrho_B)^2]^{-\frac{1}{2}} \}.$$

This expression behaves correctly in the limits  $R_{AB} \rightarrow \infty$  and  $R_{AB} \rightarrow 0$ .  $A_A$  is an atomic parameter which may be either treated as such or calculated theoretically from

$$A_A = \frac{1}{2} |\int s_A r_A \rho_A dr|.$$

#### 4. Concluding Remarks

The approach employed in the present work differs somewhat from that followed by Pople and co-workers. These latter authors were concerned about the form of the individual charge distributions in the integral (1). Here, on the other hand, attention has been focussed on the types of interactions *between* the charge distributions. The emphasis has thus been shifted from attempts to approximate charge distributions to attempts to approximate integrals.

The second member of the series of approximations proposed above is the scheme in which dipole-dipole and monopole-quadrupole type interactions are included. These both decrease asymptotically proportionally to the inverse third power of the distance. Of the additional terms included the dipole-dipole type will probably be the most important, since no cancellation between electron-electron and electron-core interactions takes place in this case.

When quadrupoles are introduced the inconvenience that charge distributions (even from an unhybridized basis set) may possess both a monopole and a quadrupole arises. Invariance with respect to rotation of local coordinate systems requires that *all* quadrupoles be retained. On the other hand, the inclusion of monopole-quadrupole terms leads to a difference between the direct interactions of *s*- and *p*-electrons with the surroundings. This difference has been asserted to be important by several authors [5, 13]. Rather complex local transformations of *d*-type functions have to be performed in this case. However, problems of this kind do not constitute major obstacles to the applicability of the approximation scheme, and it appears that the multipole-multipole decomposition of integrals is a convenient way of approaching the NDDO level of approximation in a manner which, by its very nature, suggests a detailed semiempirical parameterization scheme.

*Acknowledgement.* The author wishes to thank Dr. Å. E. Hansen for his stimulating interest in the present work.

### References

1. Pople, J. A., Santry, D. P., Segal, G. A.: *J. Chem. Phys.* **43**, S 129 (1965)
2. Pople, J. A., Beveridge, D. L., Dobosh, P. A.: *J. Chem. Phys.* **47**, 2026 (1967)
3. Pople, J. A., Beveridge, D. L.: *Approximate molecular orbital theory*. New York: McGraw-Hill 1970
4. Murrell, J. N., Harget, A. J.: *Semi-empirical self-consistent-field molecular-orbital theory of molecules*. London: Wiley 1972
5. Dewar, M. S. J., Klopman, G.: *J. Am. Chem. Soc.* **89**, 3089 (1967)
6. Jackson, J. D.: *Classical electrodynamics*, (a) Sec. 3.5. (b) Sec. 4.1. New York: Wiley 1962
7. Dahl, J. P.: *Acta Chem. Scand.* **21**, 1244 (1967)
8. Rose, M. E.: *Elementary theory of angular momentum*, Sec. 14. New York: Wiley 1957
9. Dixon, R. N.: *Mol. Phys.* **12**, 83 (1967)
10. Roothaan, C. C. J.: *J. Chem. Phys.* **19**, 1445 (1951)
11. Ohno, K.: *Theoret. Chim. Acta (Berl.)* **2**, 219 (1964)
12. Klopman, G.: *J. Am. Chem. Soc.* **86**, 4550 (1964)
13. Nicholson, B. J.: *Advan. Chem. Phys.* **18**, 249 (1970)

Dr. B. Voigt  
Department of Physical Chemistry  
H. C. Ørsted Institute  
University of Copenhagen  
DK-2100 Copenhagen Ø  
Denmark



# A Theory of Molecules in Molecules

## III. Application to the Hydrogen Bonding Interaction of Two FH Molecules

W. von Niessen

Lehrstuhl für Theoretische Chemie der Technischen Universität München, 8000 München 2, Germany

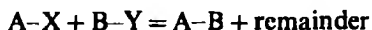
Received May 24, 1973

The theory of molecules in molecules introduced in previous articles is applied to study the hydrogen bonding interaction in the linear configuration of the dimer of FH. The transfer of localized molecular orbitals as well as the majority of the additional approximations introduced to save computational time can be justified and shown to lead to results in good agreement with those of *ab initio* calculations. An energy analysis of the effect of the hydrogen bond formation on the localized orbitals is given. It is seen that the effect is small, the major contribution to the binding energy is given by a first order perturbation treatment.

**Key words:** Transferability of localized orbitals – Molecules in molecules – FH ... FH hydrogen bonding interaction – Energy analysis

### 1. Introduction

In two previous articles a theory of molecules in molecules has been developed which permits the construction of the wave function of a molecule from the wave functions of fragment molecules, [1, 2]. (These papers will be referred to as I and II). The fundamental concept in this theory is the concept of the localized molecular orbital. Localized molecular orbitals (LMO's) are known to have the property of approximate transferability among structurally related molecules [1–14], a fact which is substantiated also by experiment [15]. In the formation of any "large" molecule A–B from molecules A–X and B–Y (where A, X, B, and Y are any molecular fragments) according to



one can distinguish in A–X and B–Y a spatial region which is only insignificantly affected by the formation of the new bonds and which – to a good approximation – can be transferred unaltered. In a one-particle approximation this means that the LMO's describing the inner shells and lone pairs of electrons and bonds in this region can be transferred from the wave functions of the fragments to the wave function of the new formed molecule. Further on there will be a spatial region – it has been called the region of interaction – where the electronic rearrangement due to the formation of the new molecule must be accounted for. The corresponding molecular orbitals have to be redetermined and cannot be transferred. The wave

function of the new molecule thus takes the form of an antisymmetrized product of LMO's transferred from the fragment wave functions and of the molecular orbitals (MO's) determined for this molecule. A projection operator is used to obtain MO's in the region of interaction which are orthogonal to the core orbitals. This condition of orthogonality is incorporated into the defining equations by the method of outer projections [16]. A similar approach has been suggested by Huzinaga and Cantu [17].

In order to give the possibility to save computational time further approximations have been introduced:

- 1) The expansion of the MO's in the region of interaction is truncated to include only those basis functions which are regarded as essential for their expansion. This subset is denoted by  $\Gamma$ . Basis functions whose centers are distant from the region of interaction will contribute only insignificantly and can be excluded from the entire set of basis functions.
- 2) The LMO's in the projection operator for orthogonality to which orthogonality can be expected because of their spatial separation from the region of interaction are taken out of the projection operator and the expansion of the remaining LMO's is restricted to the same subset  $\Gamma$  of basis functions mentioned above.
- 3) The nonorthogonality of the MO's is neglected.
- 4) The Coulomb integrals between a LMO transferred for one of the fragments and a LMO transferred for the other one are calculated by a point charge approximation and the corresponding exchange integrals are neglected. These approximations lead to a reduction of the dimension of the matrices to be diagonalized and to the neglect of a part of the basic integrals.

In the subsequent section the method of molecules in molecules (MIM) is applied in various forms to study the hydrogen bonding interaction of two FH molecules. The results are compared with more exact SCF results.

## 2. Application to FH-FH

In I the interaction of two FH molecules has been examined along the coordinate  $\text{FH} \cdots \text{HF}$  which yields a repulsive potential curve. In this article the theory of molecules in molecules is applied to study the hydrogen bonding interaction along the coordinate FH-FH. The investigations serve two purposes. The approximations introduced in I and II and briefly described in the previous section have to be justified and to be examined for their range of validity. The theory of molecules in molecules can give information on the energy contributions of the individual inner shell, lone pair, and bond orbitals to the total energy change of a process in a relatively nonarbitrary way. This question which is of particular interest to chemists will be investigated as well. The experimentally observed structure of  $(\text{FH})_2$  is bent with an angle of  $140^\circ \pm 5^\circ$  between the two FH molecules [18]. The linear as well as the bent geometry have been theoretically treated e.g. by Diercksen and Kraemer using a large basis set of Cartesian Gaussian functions [19]. The potential curve for the angular variation is very shallow having energy variations of approximately  $10^{-4}$  a.u. and can be reliably calculated if at all only with a large basis. In this article the hydrogen bonding interaction will be examined only for the linear configuration of  $(\text{FH})_2$ .

Table 1. Total SCF energies for the linear configuration of the hydrogen bonded dimer of FH. The bond distance of the two FH molecules is kept fixed at its experimental value of 1.7328 a.u. (all values in atomic units)

$R_{FF}$	$E^{SCF}$
4.5	-200.043665
4.75	-200.047530
5.0	-200.048932
5.25	-200.049255
5.5	-200.048825
6.0	-200.047361
8.0	-200.042473
13.0	-200.039773
$\infty$	-200.038991

The basis set used in the present calculations consists of 9s-type [20] and 5p-type [21] Gaussian lobe functions on the F atoms contracted to 5s-type [20] and 3p-type [21] functions and of 4s-type functions on the H atoms contracted to 3s-type functions [20]. The basis set differs slightly from the one used in I and gives a lower energy of  $E^{SCF} = -100.019495$  a.u. for the experimental bond length of  $R = 1.7328$  a.u. This compares with the best value reported in the literature of  $E^{SCF} = -100.07046$  a.u., which is believed to represent the Hartree-Fock (HF) limit [22]. The bond length of the two FH molecules is held fixed at its experimental value in all calculations, only the distance between the two molecules is varied for the linear structure in the range from  $R_{FF} = 4.5$  to 13.0 a.u. The total SCF energies are given in Table 1 and the corresponding potential curve is plotted in Fig. 1, curve a. The SCF calculations result in a binding energy of  $B = 6.4$  kcal/mole and a bond length of  $R_{FF} = 5.25$  a.u. The more extensive calculations of Diercksen *et al.* give a binding energy of  $B = 4.5$  kcal/mole at a bond length of  $R_{FF} = 5.5$  a.u. [19]. The experimental result is  $B_{exp} = 6.0$  kcal/mole [18], the bond distance being known only for the cyclic hexamer ( $R_F$ , 'hexamer') = 4.85 a.u. [23]).

The theory of molecules in molecules has been applied to this system in a number of approximate forms which will be described in the sequence of decreasing accuracy. In order to be able to transfer LMO's from the wave functions of the separate FH molecules to the dimer the canonical MO's of the SCF calculation on FH are localized using the method of Boys [24]. For the proton donor molecule  $F_1H_1$  all LMO's are transferred except the FH bond orbital, which is re-determined. For the proton acceptor FH molecule ( $F_2H_2$ ) the three lone pair orbitals are recalculated, the F inner shell and the FH bond orbital are transferred. If all basis functions are included in the set  $\Gamma$  (there are altogether 34 contracted functions in both molecules taken together) this approximation is denoted by 4, 2, 4  $\Gamma$  34. In this notation is given the number of LMO's transferred for molecule  $F_1H_1$  (4), for molecule  $F_2H_2$  (2), the number of MO's to be determined in the region of interaction (4) and the number of basis functions included in the set  $\Gamma$ . If the  $\pi$ -type basis functions on the fluorine atom  $F_1$  are excluded from the expansion of the MO's in the region of interaction, the approximation 4, 2, 4  $\Gamma$  28 is obtained. Since the three lone pair LMO's on the fluorine atom  $F_1$  are identical in the remain-

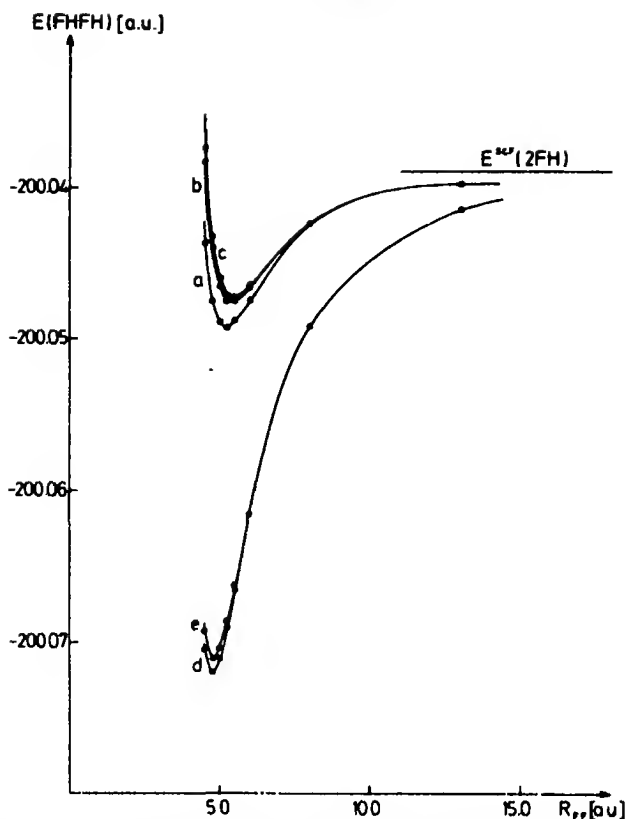


Fig. 1. Potential energy curve for FH·FH, approximation 4, 2, 4  $\Gamma$  34 (for notation and definition of *a*, *b*, *c*, *d*, *e* see text)

ing  $\sigma$ -type basis functions the orthogonality conditions on the orbitals in the interaction region can be satisfied by including only one lone pair LMO in the projection operator for orthogonality. In the approximation denoted by 4, 2, 4  $\Gamma$  26 the contracted  $s$ -type basis functions on the atoms  $F_1$  and  $H_2$  are taken out as well as as the  $\pi$ -type basis functions on atom  $F_1$ . If in addition all remaining basis functions are removed from atom  $H_2$  (this is the atom which is at the greatest distance from the region of interaction as defined above) the approximation 4, 2, 4  $\Gamma$  24 is obtained. In the crudest approximation which does not allow for any electronic rearrangement all MO's are transferred: 5, 5, 0  $\Gamma$  34. This corresponds to a first order perturbation treatment.

The results for the approximation 4, 2, 4  $\Gamma$  34 are plotted in Fig. 1 (curves *b*, *c*, *d*, and *e*) together with the exactly calculated SCF potential curve *a*. The letters in Fig. 1 have the following meaning: *b*: energy value calculated exactly, nonorthogonality of the MO's taken into account [25], *c*: as *b* only the nonorthogonality is neglected; *d* and *e* correspond to *b* and *c* but involve the point charge approximation in the calculation of the interaction energy between the two sets of transferred LMO's. Curves *b* and *c* are quite good approximations to the SCF result giving a



Table 2. Binding energies  $B$  and bond distances  $R_{FF}$  for the hydrogen bonded dimer of FH. In all cases the energy value at  $R_{FF} = 13$  a. u. has been taken as reference value for infinite separation of the two molecules. (For the notation see text)

Method	$B$ [kcal/mole]	$R_{FF}$ [a. u.]
SCF	6.0	5.25
4, 2, 4 $\Gamma$ 34 (b)	4.9	5.5
4, 2, 4 $\Gamma$ 34 (c)	4.8	5.5
4, 2, 4 $\Gamma$ 34 (d)	19.2	4.75
4, 2, 4 $\Gamma$ 34 (e)	18.6	4.75
4, 2, 4 $\Gamma$ 28 (b)	4.9	5.5
4, 2, 4 $\Gamma$ 28 (c)	4.8	5.5
4, 2, 4 $\Gamma$ 28 (d)	19.2	4.75
4, 2, 4 $\Gamma$ 28 (e)	18.9	4.75
4, 2, 4 $\Gamma$ 26 (b)	4.4	5.5
4, 2, 4 $\Gamma$ 26 (c)	4.5	5.5
4, 2, 4 $\Gamma$ 26 (d)	18.4	4.75
4, 2, 4 $\Gamma$ 26 (e)	18.2	4.75
4, 2, 4 $\Gamma$ 24 (b)	5.1	5.25
4, 2, 4 $\Gamma$ 24 (c)	3.7	5.5
4, 2, 4 $\Gamma$ 24 (d)	22.2	4.75
4, 2, 4 $\Gamma$ 24 (e)	15.9	5.0
5, 5, 0 $\Gamma$ 34 (b)	4.4	5.5

binding energy of  $B = 4.9$  kcal/mole (b) and  $B = 4.8$  kcal/mole (c) and a bond length of  $R_{FF} = 5.5$  a. u. (b and c). The SCF results are  $B = 6.0$  kcal/mole and  $R_{FF} = 5.25$  a. u. In all cases the binding energy has been calculated as the difference of the energy value at  $R_{FF} = 13.0$  a. u. and at the energy minimum. This has been done because for the other approximations the energy value at  $R_{FF} = 13.0$  a. u. has to be taken as the reference value for the following reason. As has been remarked in II the truncation of the basis set results in a nonorthogonality of the MO's calculated for one of the fragments to the transferred LMO's of the same fragment even at infinite separation of the two fragments. This constant part of the orthogonality error has the consequence that the energy value in the theory of molecules in molecules,  $E^{MIM}$ , does not approach the value  $E^{SCF}$  for  $R \rightarrow \infty$  and further on the energy values  $E^{MIM}(b)$  and  $E^{MIM}(c)$ , i. e. with the nonorthogonality taken into account or neglected, respectively, do not go to the same limiting value. Only the differences  $E^{MIM} - E^{SCF}$  and  $E^{MIM}(b) - E^{MIM}(c)$  become constant for  $R \rightarrow \infty$  and a parallel shift of the potential curves is obtained. The approximations 4, 2, 4  $\Gamma$  34 d and e reproduce astonishingly well the SCF value of the bond length:  $R_{FF} = 4.75$  a. u. (for d and e) but the binding energy is too large by a factor of about three:  $B = 19.2$  kcal/mole (d) and  $B = 18.6$  kcal/mole (e). The binding energies and bond distances for all approximations investigated are summarized in Table 2.

The energy values calculated in the approximation 4, 2, 4  $\Gamma$  28 agree extremely well with the results of approximation 4, 2, 4  $\Gamma$  34. The potential curves are not plotted because in the scale employed they would be indistinguishable from the curves in Fig. 1, which can thus be taken to represent this result as well. This ap-

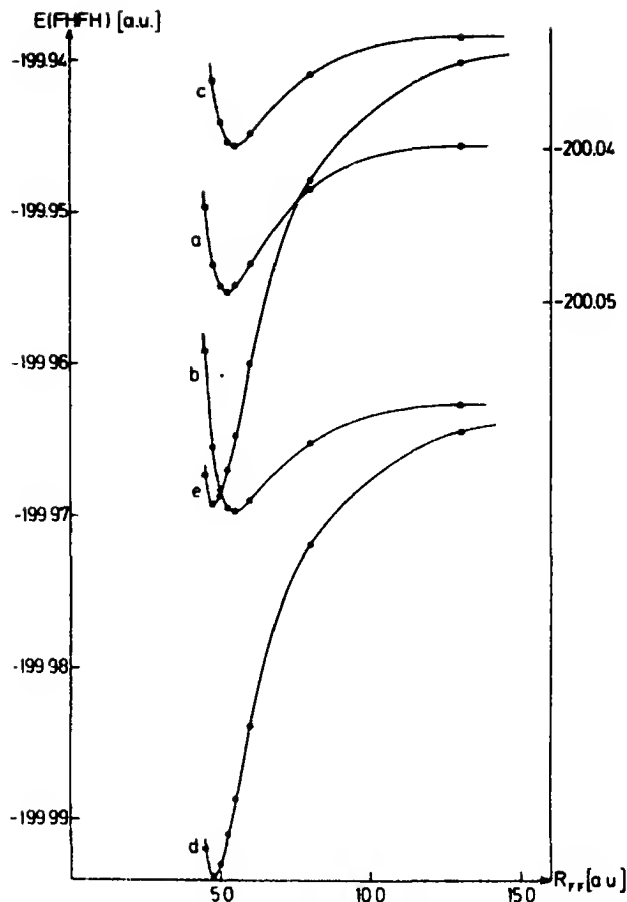


Fig. 2 Potential energy curve for FH...FH, approximation 4, 2, 4  $\Gamma$  26. (For notation and definition of *a*, *b*, *c*, *d*, *e* see text, *a* refers to the right scale, *b*, *c*, *d*, and *e* refer to the left scale)

proximation does not introduce any additional error which is indicative of the fact that the two sets of lone pair orbitals on the fluorine atoms  $F_1$  and  $F_2$  are well separated and a modification of the lone pair orbitals on the atom  $F_2$  is mainly restricted to the  $\sigma$ -type basis functions of their expansion.

For the approximation denoted by 4, 2, 4  $\Gamma$  26 the potential curves are plotted in Fig. 2. The letters *a*, *b*, *c*, *d*, and *e* have the same meaning as described above. The curves show a wider spread because of the effects on the limiting value of the total energy introduced by the truncation of the basis set. But because results *b* and *d* and results *c* and *e* involve the same approximations (except for the point charge approximation which becomes exact at infinite separation of the two molecules) they should go to the same limit as  $R \rightarrow \infty$ , which is seen to be the case. The binding energies and bond lengths calculated in this approximation compare quite well with the results of the approximation 4, 2, 4  $\Gamma$  34 (Table 2). Also the potential curves on the whole show a very similar behaviour. This is satisfactory.

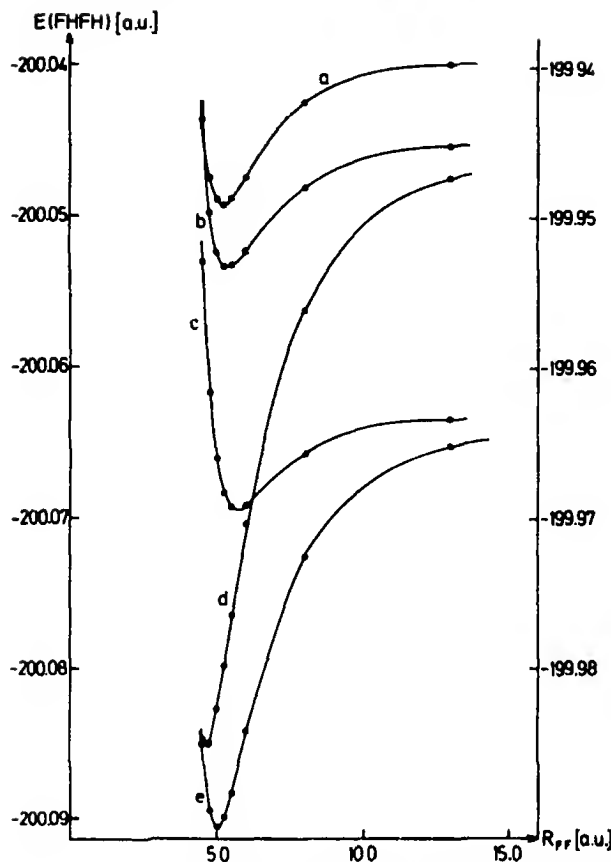


Fig. 3. Potential energy curve for FH-FH, approximation 4, 2, 4  $\Gamma$  24. (For notation and definition of *a*, *b*, *c*, *d*, *e* see text; *a*, *c*, *e* refer to the left scale, *b* and *d* to the right scale)

The results for the approximation 4, 2, 4  $\Gamma$  24 (Fig. 3, *a*, *b*, *c*, *d*, *e* as defined above) do not introduce new features compared to the previous one and can therefore be described in the same way. The approximation is poorer demonstrated by the fact that the potential curves deviate stronger from a parallel shift of the SCF potential curve or the curves of approximation 4, 2, 4  $\Gamma$  34. The binding energies and bond lengths in Table 2 consequently show a greater deviation from the values of the other approximations and results *b* and *c* as well as *d* and *e* differ more substantially than in the previous cases where the agreement can be considered to be very good. This indicates that the approximations are justified and it is first the case 4, 2, 4  $\Gamma$  24 which leads to a departure from this agreement.

The last case to be considered is the approximation 5, 5, 0  $\Gamma$  34 in which all LMO's have been transferred (Fig. 4, *a*, *b*, *c*, *d*, *e* as defined above). Only the potential curve corresponding to the exact calculation of the energy with the nonorthogonality of the MO's taken into account has been plotted in Fig. 4. It is the only reasonable result and it is astonishingly good. This first order perturbation treat-

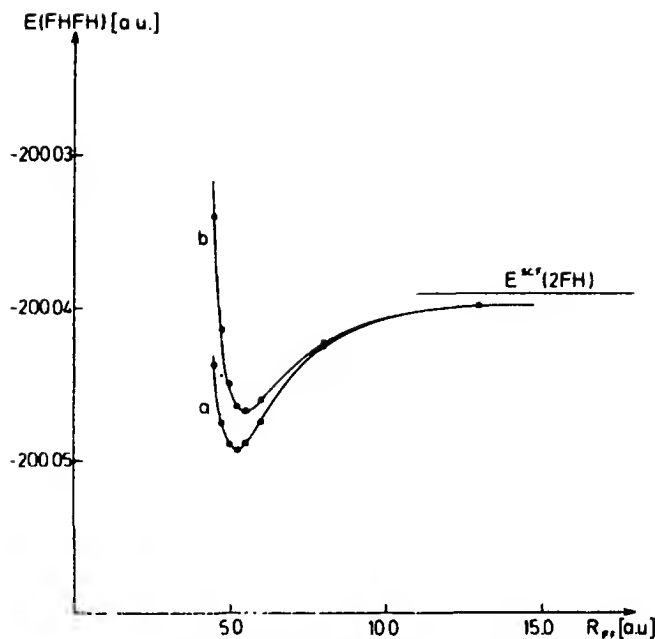


Fig. 4. Potential energy curve for FH...FH, approximation 5, 5, 0  $\Gamma$  34 (For notation and definition of  $a$  and  $b$  see text)

ment gives a binding energy of  $B = 4.4$  kcal/mole and a bond length of  $R_{FF} = 5.5$  a.u. which compare quite well with the SCF values in Table 2. The nonorthogonality of the MO's cannot be neglected in the neighbourhood of the energy minimum, nor can the point charge approximation be expected to work.

Comparing the results of approximations 4, 2, 4  $\Gamma$  34 and 5, 5, 0  $\Gamma$  34 in Figs. 1 and 4 one arrives at the conclusion that the major improvement beyond a first order perturbation calculation has not been achieved by the approximation 4, 2, 4  $\Gamma$  34. Since the effect of the hydrogen bond formation cannot be important for the inner shell orbitals (this will be substantiated in the next paragraph) it must be the lone pair orbitals on the fluorine atom  $F_1$  and the FH bond orbital in the molecule  $F_2H_2$  which account for the larger part of the remaining energy change contrary to what one would assume. The energy change arising from a modification of these orbitals must be larger than the energy change due to the recalculation of the other lone pair and bond orbitals (bond orbital in  $F_1H_1$  and lone pair orbitals in  $F_2H_2$ ). In the energy analysis discussed in the next paragraph this will be seen too. An explanation might be that the charge transfer connected with the formation of the hydrogen bond is mainly a charge migration from the bond orbital in the proton acceptor molecule  $F_2H_2$  to the proton donor molecule. The fluorine atom  $F_2$  will serve only as a charge transmitter not donating charge itself because of its high electronegativity. This point will not be examined further here. The theory of molecules in molecules could have been applied in two ways which are compatible with the requirement that the neglect of the nonorthogonality of the MO's should be a justified approximation. One path has been followed in the

present work. The second possibility would be to transfer all LMO's of the proton donor molecule and to recalculate the MO's of the proton acceptor molecules except for the inner shell MO. This would have given a better approximation to the SCF potential curve than the one of case 4, 2, 4  $\Gamma$  34. But little further information would have been derived and it was not regarded worthwhile to calculate this additional potential curve.

### 3. Energy Analysis

In the beginning of this section it has been mentioned that this theory of molecules in molecules is capable of giving information on the energy contributions of the individual LMO's to the total energy change of a process. To formulate this problem precisely it must be stated in the following form: Given any LMO in a fragment molecule, what is the contribution to the total energy if this LMO is recalculated in the region of interaction compared to a transfer of the unmodified LMO, i.e. what is the energy change beyond a first order perturbation treatment. The answer depends to a certain degree on the order in which the LMO's are included in the region of interaction, but because of the localized character of the transferred orbitals the result cannot be altogether arbitrary.

For the purpose of this energy analysis additional calculations have been performed at the theoretically determined bond distance  $R_{FF} = 5.25$  a.u. In these calculations all basis functions have been included in the set  $\Gamma$ . In the first of these additional calculations denoted by (4, 5, 1) only the FH bond orbital in the molecule  $F_1H_1$  is redetermined, all other MO's are transferred. In the calculation denoted by (5, 2, 3) only the three lone pair orbitals on the fluorine atom  $F_2$  are recalculated. In approximation (4, 1, 5) the FH bond orbital in  $F_1H_1$  is redetermined and for molecule  $F_2H_2$  all LMO's except the inner shell orbital. In approximation (1, 2, 7) only the inner shell orbitals and the FH bond orbital in the molecule  $F_2H_2$  are transferred. (1, 1, 8) denotes the approximation in which only the two inner shell orbitals are transferred. Approximations (0, 1, 9) and (1, 0, 9) finally serve to estimate the effect of the hydrogen bond formation on the inner shell MO on atom  $F_2$ : (0, 1, 9) (the inner shell on atom  $F_2$  is transferred) and in the other case on the inner shell MO on atom  $F_1$ : (1, 0, 9). These calculations are not complete to answer any question, but the main questions of interest to a chemist can be answered. The energy expectation values have been collected in Table 3 and the results of approximation 4, 2, 4  $\Gamma$  34 = (4, 2, 4) and 5, 5, 0  $\Gamma$  34 = (5, 5, 0) have been added. The letters *b*, *c*, *d*, and *e* have the same meaning as defined above. The analysis is given for the exact calculation of the energy with the nonorthogonality of the MO's taken into account (*b*), the analysis for the other cases *c*, *d*, and *e* cannot be given because the approximation (5, 5, 0) does not supply a reasonable reference. In all cases the reference value for the binding energy calculation is the energy value at infinite separation of the two molecules. From Table 3 one can extract the following information, which is summarized in Table 4.

1) At the distance  $R_{FF} = 5.25$  a.u. a first order perturbation calculation gives already a binding energy of  $B = 4.6$  kcal/mole, which is equal to 77% of the energy decrease due to the formation of the hydrogen bond (6.4 kcal/mole).

Table 3. Total energies for the hydrogen bonded dimer of FH in the SCF and in several forms of the MIM approximation.  $R_{FF}$  is kept fixed at the theoretical minimum energy distance of  $R_{FF} = 5.25$  a.u. (For notation see text; all values in atomic units)

Method	$E^{MIM}(b)$	$E^{MIM}(c)$	$E^{MIM}(d)$	$E^{MIM}(e)$
(5, 5, 0)	-200.046393	-200.055430	-200.111568	-200.1162
(4, 5, 1)	-200.046699	-200.046719	-200.031426	-200.0310
(5, 2, 3)	-200.047184	-200.047000	-200.108410	-200.1074
(4, 2, 4)	-200.047494	-200.047204	-200.069063	-200.0686
(4, 1, 5)	-200.048729	-200.048672	-200.043227	-200.0431
(1, 2, 7)	-200.047877	-200.047876	-200.056233	-200.0562
(1, 1, 8)	-200.049142	-200.049142	-200.049153	-200.0491
(1, 0, 9)	-200.049233	-200.049233	—	—
(0, 1, 9)	-200.049164	-200.049164	—	—
SCF = (0, 0, 10)	-200.049255	—	—	—

Table 4. Contributions to the total hydrogen bonding energy of two FH molecules (at the distance  $R_{FF} = 5.25$  a.u.) obtained by modifying the LMO's in the proton donor molecule  $F_1H_1$  and in the proton acceptor molecule  $F_2H_2$  (all values in kcal/mole)

Modified LMO	$\Delta E$
None	4.6
Lone pair LMO's in $F_1H_1$	0.24
Bond LMO in $F_1H_1$	0.19
Lone pair LMO's in $F_2H_2$	0.5
Bond LMO in $F_2H_2$	0.78
Inner shell LMO in $F_1H_1$	0.014
Inner shell LMO in $F_2H_2$	0.056
	6.38

2) The modification of only the FH bond orbital in the proton donor molecule gives beyond a first order perturbation treatment an energy lowering 0.19 kcal/mole.

3) Modifying only the three lone pair orbitals of the proton acceptor molecule yields an energy lowering of 0.5 kcal/mole or 0.17 kcal/mole per lone pair LMO. It is seen that the effect of the hydrogen bond formation on the lone pair orbitals of the proton acceptor molecule is greater than its effect on the FH bond orbital of the proton donor molecule by 0.31 kcal/mole.

4) If the FH bond orbital in the proton donor molecule and the three lone pair orbitals in the proton acceptor molecule are modified together, the energy lowering is 0.69 kcal/mole, which is within  $10^{-6}$  a.u. the sum of the data given in 2) and 3). Localized molecular orbitals thus describe each separated region of the molecule.

5) If in addition to 4) the FH bond orbital in the proton acceptor molecule is redetermined a further energy lowering of 0.78 kcal/mole results.

6) If in addition to 4) the lone pair orbitals in the proton donor molecule are recalculated one obtains compared to 4) an energy lowering of 0.24 kcal/mole.

0.08 kcal/mole per lone pair orbital. The effect of the hydrogen bond formation is consequently greater on the FH bond orbital of the proton acceptor molecule than on the lone pair orbitals of the proton donor molecule by 0.54 kcal/mole. It is somewhat surprising to see that the sum of 5) and 6) is a larger energy contribution than the sum of 2) and 3) and that the modification of the bond orbital in  $F_2H_2$  is energetically most important. This point has been discussed above.

7) If only the inner shell LMO's are transferred and all other MO's recalculated the energy lowering compared to 4) is 1.03 kcal/mole. The energy lowering of 5) and 6) are thus additive to within  $10^{-5}$  a.u. which should be the case because of the separation of the two molecular regions described by the FH bond orbital of the proton acceptor molecule and the lone pair orbitals of the proton donor molecule.

8) When all MO's are redetermined except the two inner shell LMO's the energy difference to the SCF result is 0.07 kcal/mole. This can be decomposed exactly into a very small contribution of 0.014 kcal/mole due to modifying the inner shell LMO of the proton donor molecule and a larger contribution of 0.056 kcal/mole due to modifying the inner shell LMO of the proton acceptor molecule. The reason for this difference is presumably that the greater energetic effects of the hydrogen bond formation on the proton acceptor molecule cause also a larger influence on the  $F_2$  inner shell LMO compared to the  $F_1$  inner shell LMO.

These data supply some valuable chemical information on hydrogen bonds involving hydrogen fluoride.

#### 4. Conclusions

The theory of molecules in molecules has been applied in various approximate forms to study the hydrogen bonding interaction between two FH molecules. Most of these approximate forms have been reasonably successful and one can draw the following conclusions from these applications. The transfer of LMO's from the fragment molecules to the molecule to be calculated has been justified again in the present case. This approach gives a satisfactory answer also in the case where all LMO's of the fragments are transferred, accounting already for about 75% of the binding energy between two FH molecules, if no further approximations are made. The modification of the LMO's in the fragments gives only small energy contributions ranging from 0.2-0.8 kcal/mole except for the inner shell LMO's, whose contributions are smaller. The energetic importance of the modifications of the LMO's decreases in the sequence: bond orbital in the proton acceptor molecule, lone pair LMO's in the proton acceptor molecule, lone pair LMO's in the proton donor molecule, bond orbital in the proton donor molecule. This order has been discussed above. The energy analysis has also shown that LMO's describe separate regions in a molecule and that energetic effects due to modifying orbitals describing different regions are additive to a high accuracy.

The truncation of the basis set for the expansion of the MO's in the region of interaction and in the projection operator for orthogonality (approximations denoted by  $\Gamma$  34,  $\Gamma$  28,  $\Gamma$  26, and  $\Gamma$  24) leads to satisfactory potential curves which do not deviate strongly from a parallelly shifted SCF potential curve and reproduce the binding energy and bond distance to a satisfactory degree. The agreement

deteriorates only for the approximation 4, 2, 4  $\Gamma$  24. A number of basis functions can thus be left out from the basis set for the expansion of the MO's in the region of interaction: This result is not of particular importance for the small molecules considered in the present investigation, but the validity of this approximation already in this case raises the expectation that it may become a key to saving computational time.

For the purpose of saving computational time it is further on necessary to neglect the nonorthogonality of the MO's, otherwise all integrals will appear again in the expression for the total energy even if they are not required for the self-consistency process in the calculation. This approximation has been justified by the calculations in the same way as the truncation of the basis set. For all approximations the potential curves reproduce satisfactorily the binding energy and bond distance of the hydrogen fluoride dimer and run parallel to the SCF potential curve to a satisfactory degree (only the approximation 4, 2, 4  $\Gamma$  24 has a somewhat poorer quality). The neglect of the nonorthogonality in the calculation of the total energy expectation value is thus valid in the same cases, for the same distances and to the same accuracy as the truncation of the basis set is valid. This was remarked already in II.

The point charge approximation for the calculation of the interaction energy between the two sets of transferred LMO's is the weakest point of the present version of the theory. But as mentioned before it can be refined in several ways; it has been chosen because it represents the simplest method possible. The approximation does not fail, however. The bond distances are fairly well reproduced as can be seen from Figs. 1 to 3, they are too small by about 10% compared to the SCF value, whereas the binding energies turn out too large by a factor of about three. The essential feature, however, is that the bonding is reproduced, even if the point charge approximation is used. In a forthcoming article [26], in which the theory of molecules in molecules is applied to study the hydrogen bonding interaction between an ammonia molecule as proton acceptor and a water molecule as proton donor, it will be seen that the point charge approximation is not applicable.

*Acknowledgment.* The calculations were performed on an IBM 360/91 computer while the author was at the Max-Planck-Institut für Physik und Astrophysik, München. The author would like to thank Professor Dr. L. Biermann and Dr. G. Diercksen for the opportunity of working in the institute and for using its facilities. Financial support of the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

## References

1. von Niessen, W.: J. Chem. Phys. **55**, 1948 (1971)
2. von Niessen, W.: Theoret. Chim. Acta (Berl.), accepted for publication
3. Edmiston, C., Ruedenberg, K.: Rev. Mod. Phys. **35**, 457 (1963); J. Chem. Phys. **43**, S. 97 (1965)
4. England, W., Salmon, L. S., Ruedenberg, K.: Topics in current chemistry, Vol. 23, p. 31. Berlin: Springer 1971
5. Weinstein, H., Pauncz, R., Cohen, M.: Advances at. mol. phys., Vol. 7, p. 97. New York: Academic 1971
6. von Niessen, W.: J. Chem. Phys. **56**, 4290 (1972); Theoret. Chim. Acta (Berl.) **27**, 9 (1972); **29**, 29 (1973)
7. Rothenberg, S.: J. Chem. Phys. **51**, 3389 (1969); J. Am. Chem. Soc. **93**, 68 (1970)
8. Switkes, F., Lipscomb, W. N., Newton, M. D.: J. Am. Chem. Soc. **92**, 3837, 3847 (1970)



9. Newton, M. D., Switkes, E.: J. Chem. Phys. **54**, 3179 (1971)
10. Levy, M., Stevens, W. J., Shull, H., Hagstrom, S.: J. Chem. Phys. **52**, 5483 (1970)
11. Adams, W. H.: J. Am. Chem. Soc. **92**, 2198 (1969); J. Chem. Phys. **57**, 2340 (1972)
12. Peters, D.: J. Am. Chem. Soc. **94**, 707 (1972)
13. England, W., Gordon, M. S.: J. Am. Chem. Soc. **93**, 4649 (1971); **94**, 4818, 5168 (1972); Chem. Phys. Letters **15**, 59 (1972)
14. Epstein, I. R.: J. Chem. Phys. **53**, 4425 (1970); Roux, M., Epstein, I. R.: Chem. Phys. Letters **18**, 18 (1973)
15. Eisenberger, P., Marra, W. C.: Phys. Rev. Letters **27**, 1413 (1971)
16. Löwdin, P. O.: Phys. Rev. **139**, A 357 (1965)
17. Huzinaga, S., Cantu, A. A.: J. Chem. Phys. **55**, 5543 (1971)
18. Smith, D. F.: J. Mol. Spectry **3**, 473 (1959)
19. Diercksen, G. H. F., Kraemer, W. P.: Chem. Phys. Letters **6**, 419 (1970)
20. Huzinaga, S.: J. Chem. Phys. **42**, 1293 (1965)
21. Whitten, J. L.: J. Chem. Phys. **44**, 359 (1966)
22. McLean, A. D., Yoshimine, M.: J. Chem. Phys. **46**, 3682 (1967); Cade, P. E., Huo, W. M.: J. Chem. Phys. **47**, 614 (1967)
23. Janzen, J., Bartell, L. S.: U.S. At. Energy Comm. IS-1940 (1968); J. Chem. Phys. **50**, 3611 (1969)
24. Boys, S. F.: Rev. Mod. Phys. **32**, 296 (1960). – Foster, J. M., Boys, S. F.: Rev. Mod. Phys. **32**, 300 (1960); Boys, S. F., In: Löwdin, P. O. (Ed.): Quantum theory of atoms, molecules and the solid state, p. 253. New York: Academic 1966
25. Löwdin, P. O.: J. Chem. Phys. **18**, 365 (1950)
26. von Niessen, W.: Theoret. chim. Acta (Berl.), in press

Dr. W. von Niessen  
Lehrstuhl für Theoretische Chemie  
der Technischen Universität München  
D-8000 München 2, Arcissstraße 21  
Federal Republic of Germany



# Point Charge Models for LiH, CH<sub>4</sub>, and H<sub>2</sub>O

A. D. Tait and G. G. Hall

Department of Mathematics, University of Nottingham, Nottingham, NG7 2RD, England

Received April 10, 1973

Point charge models for LiH, CH<sub>4</sub>, and H<sub>2</sub>O are presented. The models preserve the correct total charge and dipole moment of the molecules. Relations between spherical Gaussian wave function values and point charge model values of a variety of one-electron molecular properties are derived. The errors inherent in some of the point charge model values are of two types: those which may be large but are easily evaluated and those which are small and diminish rapidly as the distance from the molecule increases. The models are shown to be a reliable means of calculating one-electron properties and possible uses of the models are suggested.

**Key words:** Point charge models – One-electron properties

## 1. Introduction

In a recent paper Hall [1] has proposed a point charge model for molecules. The model preserves the correct total charge of the molecule and the dipole moment. The model also gives a good approximation to the electrostatic potential outside of the shell of the molecule. In this paper we examine the point charge model as a means of calculating a variety of one-electron properties. As examples the results for LiH, CH<sub>4</sub>, and H<sub>2</sub>O are presented. Lithium hydride was chosen as a very simple ionic system, methane because of its high symmetry, and water because of the presence of lone pairs.

The electron density in a closed shell molecule is defined by

$$\rho(r) = 2 \sum_{st} p_{st} \varphi_s(r) \varphi_t(r), \quad (1)$$

where  $p_{st}$  is an element of the density matrix and is calculated here from an SCF wave function. The factor of two appears as a result of summing over the two spins.  $\varphi_s(r)$  is an orbital with centre  $r_s$ , exponent  $\alpha_s$  and normalized to 1. If  $\varphi_s$  and  $\varphi_t$  are spherical Gaussian orbitals then their product is a spherical Gaussian with centre at

$$r_{st} = (\alpha_s r_s + \alpha_t r_t) / (\alpha_s + \alpha_t),$$

and exponent

$$\alpha_{st} = (\alpha_s + \alpha_t).$$

Thus  $\rho$  is a sum of spherical Gaussians. The basis of the point charge model is an approximation to  $\rho(r)$ , namely

$$\rho^*(r) = 2 \sum_{st} p_{st} S_{st} \delta(r - r_{st}), \quad (2)$$

where  $S_{st}$  is the overlap integral. The total number of electrons  $N$ , is found by integrating  $\varrho$  over all space:

$$N = \int \varrho(\mathbf{r}) d\mathbf{r} = 2 \sum_{st} p_{st} \int \varphi_s(\mathbf{r}) \varphi_t(\mathbf{r}) d\mathbf{r} = \int \varrho^*(\mathbf{r}) d\mathbf{r}.$$

The electronic dipole moment is

$$\mu = \int \mathbf{r} \varrho(\mathbf{r}) d\mathbf{r} = 2 \sum_{st} p_{st} S_{st} \mathbf{r}_{st} = \int \mathbf{r} \varrho^*(\mathbf{r}) d\mathbf{r}.$$

The total charge is thus divided into charges of known magnitude and position. These charges satisfy the essential conservation laws that the values of  $N$  and  $\mu$  are preserved.

For the purpose of the following discussion we assume a set of point charges  $Z_{st} = 2p_{st}S_{st}$ . In practice since  $Z_{st} = Z_{ts}$  and  $\mathbf{r}_{st} = \mathbf{r}_{ts}$  the number of charges in the model can be reduced.

## 2. The Point Charge Model and One-Electron Properties

When discussing one-electron properties it is convenient to select as origin the point, at which the properties are evaluated. For a system of point charges  $Z_k$  situated at points  $(x_k, y_k, z_k)$ , the following properties may be defined:

Multiple moments

$$M = \sum_k Z_k x_k^u y_k^v z_k^w, \quad (3)$$

Potential

$$V = \sum_k' Z_k / |r_k|, \quad (4)$$

Electric field

$$E_\lambda = - \sum_k' Z_k \lambda_k / |r_k|^3, \quad (5)$$

Electric field gradient

$$q_{\lambda\sigma} = - \sum_k' Z_k (3\lambda_k \sigma_k - \delta_{\lambda\sigma} r_k^2) / |r_k|^5. \quad (6)$$

In (4) (6)  $r_k^2 = (x_k^2 + y_k^2 + z_k^2)$  and the primed summation indicates that any term with  $r_k = 0$  is omitted.  $\lambda$  or  $\sigma$  in (5) and (6) may each be  $x$ ,  $y$ , or  $z$ . The dipole, second and third moments are obtained from (3) by setting  $u + v + w$  to 1, 2, and 3 respectively ( $u$ ,  $v$ , and  $w$  are integers).

The general expression for the electronic multipole moment at the origin is

$$M_e = 2 \sum_{st} p_{st} \langle \varphi_s | x^u y^v z^w | \varphi_t \rangle. \quad (7)$$

If  $\varphi_s$  and  $\varphi_t$  are spherical Gaussian orbitals we may rewrite (7) using standard formulae [2] as

$$M_e = 2 \sum_{st} p_{st} S_{st} \sum_{i=0}^{[u/2]} \binom{u}{2i} x_{st}^{u-2i} \frac{(2i-1)!!}{(2\alpha_{st})^i} \cdot \sum_{j=0}^{[v/2]} \binom{v}{2j} y_{st}^{v-2j} \frac{(2j-1)!!}{(2\alpha_{st})^j} \sum_{k=0}^{[w/2]} \binom{w}{2k} z_{st}^{w-2k} \frac{(2k-1)!!}{(2\alpha_{st})^k}. \quad (8)$$

In (8)  $[u/2]$  is the largest integer  $\leq u/2$ , and  $x_{st}$ ,  $y_{st}$ ,  $z_{st}$  are the components of  $\mathbf{r}_{st}$ . From this latter equation we see that the components of the dipole moment have

the form

$$\mu_x = 2 \sum_{st} p_{st} S_{st} x_{st} = \sum_{st} Z_{st} x_{st} = \mu_x^* . \quad (9)$$

We will use symbols superscripted with an asterisk to represent point charge model values. The diagonal elements of the second moment tensor are of the same form as  $Q_{xx}$ , namely

$$\begin{aligned} Q_{xx} &= 2 \sum_{st} p_{st} S_{st} \left( x_{st}^2 + \frac{1}{2\alpha_{st}} \right) \\ &= \sum_{st} Z_{st} x_{st}^2 + \frac{1}{2} \sum_{st} Z_{st} / \alpha_{st} \\ &= Q_{xx}^* + \frac{1}{2} \sum_{st} Z_{st} / \alpha_{st} . \end{aligned} \quad (10)$$

The off-diagonal elements of this tensor are of the form

$$Q_{xy} = 2 \sum_{st} p_{st} S_{st} x_{st} y_{st} = \sum_{st} Z_{st} x_{st} y_{st} = Q_{xy}^* . \quad (11)$$

The third moments are of three types of which  $R_{xxx}$ ,  $R_{xyy}$  and  $R_{xyz}$  will serve as examples.

$$\begin{aligned} R_{xxx} &= 2 \sum_{st} p_{st} S_{st} (x_{st}^3 + 3 x_{st} / (2\alpha_{st})) \\ &= \sum_{st} Z_{st} x_{st}^3 + \frac{3}{2} \sum_{st} Z_{st} x_{st} / \alpha_{st} \\ &= R_{xxx}^* + \frac{3}{2} \sum_{st} Z_{st} x_{st} / \alpha_{st} \end{aligned} \quad (12)$$

$$\begin{aligned} R_{xyy} &= 2 \sum_{st} p_{st} S_{st} (y_{st}^2 + 1/2\alpha_{st}) x_{st} \\ &= \sum_{st} Z_{st} y_{st}^2 x_{st} + \frac{1}{2} \sum_{st} Z_{st} x_{st} / \alpha_{st} \\ &= R_{xyy}^* + \frac{1}{2} \sum_{st} Z_{st} x_{st} / \alpha_{st} . \end{aligned} \quad (13)$$

$$R_{xyz} = 2 \sum_{st} p_{st} S_{st} x_{st} y_{st} z_{st} = \sum_{st} Z_{st} x_{st} y_{st} z_{st} = R_{xyz}^* . \quad (14)$$

The formulae for the potential, electric fields and electric field gradients involve the function  $F_m(a)$  defined as follows

$$F_m(a) = \int_0^1 u^{2m} \exp(-au^2) du .$$

Useful related functions are

$$\bar{F}_m(a) = \int_1^\infty u^{2m} \exp(-au^2) du ,$$

and

$$\int_0^\infty u^{2m} \exp(-au^2) du = \frac{(2m-1)!!}{2(2a)^m} \left( \frac{\pi}{a} \right)^{\frac{1}{2}} .$$

Thus

$$F_m(a) = \frac{(2m-1)!!}{2(2a)^m} \left( \frac{\pi}{a} \right)^{\frac{1}{2}} - \bar{F}_m(a) .$$

For large values of  $a$   $\bar{F}_m(a)$  is approximately  $\exp(-a)/2a$ .

The electronic part of the potential is

$$\begin{aligned}
 V &= 2 \sum_{st} p_{st} \left\langle \varphi_s \left| \frac{1}{r} \right| \varphi_t \right\rangle \\
 &= 2 \sum_{st} p_{st} 2S_{st} \left( \frac{\alpha_{st}}{\pi} \right)^{\frac{1}{2}} F_0(\alpha_{st} r_{st}^2) \\
 &= 2 \sum_{st} \frac{p_{st} S_{st}}{|r_{st}|} - 4 \sum_{st} p_{st} S_{st} \left( \frac{\alpha_{st}}{\pi} \right)^{\frac{1}{2}} \bar{F}_0(\alpha_{st} r_{st}^2) \\
 &= \sum_{st} Z_{st}/|r_{st}| - \frac{2}{\sqrt{\pi}} \sum_{st} Z_{st} \sqrt{\alpha_{st}} \bar{F}_0(\alpha_{st} r_{st}^2) \\
 &= V^* - \frac{2}{\sqrt{\pi}} \sum_{st} Z_{st} \sqrt{\alpha_{st}} \bar{F}_0(\alpha_{st} r_{st}^2). \quad (15)
 \end{aligned}$$

The  $x$  component of the electric field is typical of the three components and is defined as

$$\begin{aligned}
 E_x &= 2 \sum_{st} p_{st} \left\langle \varphi_s \left| \frac{x}{r^3} \right| \varphi_t \right\rangle \\
 &= 2 \sum_{st} p_{st} S_{st} 4\alpha_{st} x_{st} \left( \frac{\alpha_{st}}{\pi} \right)^{\frac{1}{2}} \bar{F}_1(\alpha_{st} r_{st}^2) \\
 &= \sum_{st} Z_{st} x_{st}/|r_{st}|^3 - \frac{4}{\sqrt{\pi}} \sum_{st} Z_{st} \alpha_{st}^{\frac{1}{2}} x_{st} \bar{F}_1(\alpha_{st} r_{st}^2) \\
 &= E_x^* - \frac{4}{\sqrt{\pi}} \sum_{st} Z_{st} x_{st} \alpha_{st}^{\frac{1}{2}} \bar{F}_1(\alpha_{st} r_{st}^2). \quad (16)
 \end{aligned}$$

Finally the electronic components of the electronic field gradient tensor are

$$\begin{aligned}
 q_{\lambda\sigma} &= 2 \sum_{st} p_{st} \left\langle \varphi_s \left| \frac{3\lambda\sigma - \delta_{\lambda\sigma} r^2}{r^5} \right| \varphi_t \right\rangle \\
 &= 2 \sum_{st} p_{st} S_{st} \frac{4}{3} \left( \frac{\alpha_{st}^3}{\pi} \right)^{\frac{1}{2}} \{6\lambda_{st} \sigma_{st} \alpha_{st} \bar{F}_2(\alpha_{st} r_{st}^2) \\
 &\quad - \delta_{\lambda\sigma} (3\bar{F}_1(\alpha_{st} r_{st}^2) - \exp(-\alpha_{st} r_{st}^2))\} \\
 &= \sum_{st} \frac{Z_{st}}{|r_{st}|^5} (3\lambda_{st} \sigma_{st} - \delta_{\lambda\sigma} r_{st}^2) \\
 &\quad - \frac{4}{3} \sum_{st} Z_{st} \alpha_{st} \left( \frac{\alpha_{st}}{\pi} \right)^{\frac{1}{2}} \{6\lambda_{st} \sigma_{st} \alpha_{st} \bar{F}_2(\alpha_{st} r_{st}^2) \\
 &\quad - \delta_{\lambda\sigma} (3\bar{F}_1(\alpha_{st} r_{st}^2) - \exp(-\alpha_{st} r_{st}^2))\}. \quad (17) \\
 &= q_{\lambda\sigma}^* - \frac{4}{3\sqrt{\pi}} \sum_{st} Z_{st} \alpha_{st}^{\frac{1}{2}} \{6\lambda_{st} \sigma_{st} \alpha_{st} \bar{F}_2(\alpha_{st} r_{st}^2) \\
 &\quad - \delta_{\lambda\sigma} (3\bar{F}_1(\alpha_{st} r_{st}^2) - \exp(-\alpha_{st} r_{st}^2))\}.
 \end{aligned}$$

In (17)  $\lambda$  and  $\sigma$  can each be  $x$ ,  $y$ , or  $z$ . Examination of (9)–(17) shows that each one-electron property can be expressed as the point charge model value plus a correction. This correction is zero for the dipole moment, off-diagonal elements of the second moment tensor and for  $R_{xyz}$ . The correction is obviously the error obtained by using only the point charge value. The error in  $Q_{\lambda\lambda}^*$  is constant. The errors in the remaining multipole moments are related in that the error in  $R_{\lambda\lambda\lambda}^*$  is exactly three times the error in  $R_{\lambda\sigma\sigma}^*$ . Naturally if the evaluation of a third moment produces zero then the point charge model also gives zero and the error is zero. In general the error in these third moments increases with distance.

Inspection of (15)–(17) shows that the error involved in the point charge model decreases exponentially at large distances. The error decreases more rapidly than the actual value. The exponential term which appears in the correction to the diagonal term  $q_{\lambda\lambda}$  is the contribution from the singularity at  $r_{\lambda\lambda} = 0$ , and cannot be represented by the point charge model.

### 3. Examples of Point Charge Models

In this section we present the results obtained from simple spherical Gaussian wave functions for LiH, CH<sub>4</sub>, and H<sub>2</sub>O. We have used the smallest basis sets compatible with quantum mechanical principles and chemical intuition. All of the values are in atomic units and in the Tables powers of 10 are represented by numbers in parenthesis. The contribution from the nuclei of each molecule is included in each property.

#### *Lithium Hydride*

The nuclear geometry for LiH is Li at (0, 0, 0) and H at (0, 0, 3.02). The basis set consists of two spherical Gaussian orbitals which give a minimum energy when placed at (0, 0, -0.008251) and (0, 0, 2.72101) with exponents 2.00085 and 0.178880 respectively. The density matrix has diagonal elements of 1.01454 and off-diagonal elements of -0.121437. The point charge model is given in Table 1. The dipole moment has a non-zero  $z$ -component of -2.47184. The value of  $Q_{xx}$  and  $Q_{yy}$  is -3.07599 whereas  $Q_{xx}^*$  and  $Q_{yy}^*$  are both zero. The value of  $Q_{zz}$  is -8.97608 and that of  $Q_{zz}^*$  is -5.90008. The error in  $Q_{xx}^*$  is thus -3.07599. These values are constant along the  $x$ - and  $y$ -axes. The values of  $Q_{xz}$  at points on the

Table 1. Point charge model for LiH. All the charges are on the  $z$ -axis

Charge	$z$
3.00000 (0)	0.00000 (0)
1.00000 (0)	3.02000 (0)
-2.02907 (0)	-8.25054 (-3)
5.81430 (-2)	2.15726 (-1)
-2.02907 (0)	2.72101 (0)

Table 2.  $Q_{xx}$  for LiH at points along the x-axis

$x$	$Q_{xx}$	$Q_{xx}^*$
0.0	0.00000 (0)	0.00000 (0)
1.2	2.96621 (0)	2.96621 (0)
2.4	5.93242 (0)	5.93242 (0)
3.6	8.89863 (0)	8.89863 (0)
4.8	1.18648 (+1)	1.18648 (+1)
6.0	1.48311 (+1)	1.48311 (+1)
7.2	1.77973 (+1)	1.77973 (+1)

Table 3. Potential at points on the x-axis of LiH

$x$	$V$	$V^*$
1.2	4.91008 (-1)	4.82258 (-1)
2.4	1.29874 (-1)	1.28673 (-1)
3.6	4.90567 (-2)	4.89960 (-2)
4.8	2.29694 (-2)	2.29683 (-2)
6.0	1.23908 (-2)	1.23907 (-2)
7.2	7.38334 (-3)	7.38334 (-3)

x-axis are presented in Table 2, and although we have given them to only six significant digits the agreement is to ten digits which, in our case, is machine accuracy. Of the third moments  $R_{yyy}$ ,  $R_{xxy}$ ,  $R_{yzz}$ , and  $R_{xyz}$  are zero. As  $R_{xxx}^*$  and  $R_{xyy}^*$  are zero  $R_{xxx}$  and  $R_{xyy}$  represent the errors in the point charge model values, the error in  $R_{xxx}^*$  being three times that in  $R_{xyy}$  and  $R_{xzz}$ . The value of the latter is in general non-zero; inspection of (13) shows that at points on the x-axis  $R_{xzz}$  has a linear dependence on  $x$ , hence it is zero at the origin. At  $x = 1.2$   $R_{xzz}$  is 10.7713 and  $R_{xzz}^*$  is 7.08010. The error in  $R_{xzz}^*$  is 3.69119 at this point.

The potential at points along the x-axis is given in Table 3. The value of  $E_y$  is zero on the x-axis and as the value of the  $y$  co-ordinate of each point charge is zero  $E_y^*$  is zero. The  $x$ - and  $z$ -components of the electric field on the x-axis are presented in Table 4. Non-zero electric field gradients are shown in Table 5.

Table 4. Electric field at points on the x-axis of LiH

$x$	$E_x$	$E_x^*$	$E_z$	$E_z^*$
1.2	6.73672 (-1)	6.55240 (-1)	8.57307 (-2)	1.05315 (-1)
2.4	1.20862 (-1)	1.18394 (-1)	5.80802 (-2)	6.08786 (-2)
3.6	3.47653 (-2)	3.45886 (-2)	3.02142 (-2)	3.03477 (-2)
4.8	1.29986 (-2)	1.29944 (-2)	1.60428 (-2)	1.60452 (-2)
6.0	5.80397 (-3)	5.80393 (-3)	9.20809 (-3)	9.20810 (-3)
7.2	2.93884 (-3)	2.93884 (-3)	5.68425 (-3)	5.68425 (-3)



Table 5. Electric field gradients for LiH at points on the x-axis

x	$q_{xx}$	$q_{xx}^*$	$q_{yy}$	$q_{yy}^*$
1.2	-1.27896 (0)	-1.20748 (0)	6.08082 (-1)	5.46033 (-1)
2.4	-1.39458 (-1)	-1.38217 (-1)	5.33014 (-2)	4.93308 (-2)
3.6	-3.15923 (-2)	-3.13441 (-2)	9.88093 (-3)	9.60795 (-3)
4.8	-9.56834 (-3)	-9.55963 (-3)	2.71412 (-3)	2.70716 (-3)
6.0	-3.56406 (-3)	-3.56412 (-3)	9.67387 (-4)	9.67322 (-4)
7.2	-1.54122 (-3)	-1.54122 (-3)	4.08172 (-4)	4.08172 (-4)
x	$q_{xx}$	$q_{xx}^*$	$q_{zz}$	$q_{zz}^*$
1.2	6.70882 (-1)	6.61451 (-1)	3.66529 (-3)	-1.44757 (-2)
2.4	8.61564 (-2)	8.88862 (-2)	-2.92345 (-2)	-3.51444 (-2)
3.6	2.17113 (-2)	2.17361 (-2)	-1.67120 (-2)	-1.71060 (-2)
4.8	6.85439 (-3)	6.85247 (-3)	-7.95850 (-3)	-7.96738 (-3)
6.0	2.59686 (-3)	2.59680 (-3)	-3.96437 (-3)	-3.96437 (-3)
7.2	1.13305 (-3)	1.13305 (-3)	-2.13352 (-3)	-2.13352 (-3)

### Methane

The geometry of the methane molecule is close to that of Arrighini, Guidotti, Maestro, Moccia, and Salvetti [3]. The protons are situated at  $(a, a, a)$ ,  $(a, -a, -a)$ ,  $(-a, a, -a)$ , and  $(-a, -a, a)$ , the value of 1.19350 for 'a' corresponds to a C-H bond length of 2.0672. The basis set consists of five spherical Gaussians; one on the carbon nucleus with an exponent of 9.30694 and one in each C-H bond at 1.23654 from the carbon nucleus. The exponents of the bond orbitals are 0.356102.

Table 6. Point charge model for methane

Charge	x	y	z
1.00000 (0)	1.19350 (0)	1.19350 (0)	1.19350 (0)
1.00000 (0)	1.19350 (0)	-1.19350 (0)	-1.19350 (0)
1.00000 (0)	-1.19350 (0)	1.19350 (0)	-1.19350 (0)
1.00000 (0)	-1.19350 (0)	-1.19350 (0)	1.19350 (0)
3.93691 (0)	0.00000 (0)	0.00000 (0)	0.00000 (0)
-3.11650 (0)	7.13916 (-1)	7.13916 (-1)	7.13916 (-1)
7.33820 (-1)	7.13916 (-1)	0.00000 (0)	0.00000 (0)
-3.11650 (0)	7.13916 (-1)	-7.13916 (-1)	-7.13916 (-1)
7.33820 (-1)	0.00000 (0)	7.13916 (-1)	0.00000 (0)
7.33820 (-1)	0.00000 (0)	0.00000 (0)	-7.13916 (-1)
-3.11650 (0)	-7.13916 (-1)	7.13716 (-1)	-7.13916 (-1)
7.33820 (-1)	0.00000 (0)	0.00000 (0)	7.13916 (-1)
7.33820 (-1)	0.00000 (0)	-7.13916 (-1)	0.00000 (0)
7.33820 (-1)	-7.13916 (-1)	0.00000 (0)	0.00000 (0)
-3.11650 (0)	-7.13916 (-1)	-7.13916 (-1)	7.13916 (-1)
3.15441 (-2)	2.63092 (-2)	2.63092 (-2)	2.63092 (-2)
3.15441 (-2)	2.63092 (-2)	-2.63092 (-2)	-2.63092 (-2)
3.15441 (-2)	-2.63092 (-2)	2.63092 (-2)	-2.63092 (-2)
3.15441 (-2)	-2.63092 (-2)	-2.63092 (-2)	2.63092 (-2)

The density matrix for  $\text{CH}_4$  has elements as follows:  $p_{ii}(i < 5) = 1.55825$ ,  $p_{55} = 1.03154$ ,  $p_{ij} = p_{ji}$  ( $i < 5, j < 5, i \neq j$ ) =  $-0.379159$ , and the remaining elements are  $-0.057604$ . The point charge model consists of the nineteen charges listed in Table 6.

As expected, the dipole moment for  $\text{CH}_4$  is zero. The diagonal second moments  $Q_{\lambda\lambda}$  are equal at  $-5.61731$  and  $Q_{\lambda\lambda}^* = 0.0922448$ .  $Q_{\lambda\lambda}$  and  $Q_{\lambda\lambda}^*$  are constant along the C-H bond direction.  $Q_{\lambda\sigma}$  is zero. There are only three distinct values for the third moments. As a consequence of the symmetry of methane  $R_{xxx} = R_{yyy} = R_{zzz}$ ,  $R_{xyz}$  is unique and the six remaining third moments are equal. Values of the third moments at points along a C-H bond are directly proportional to the distance from the carbon nucleus, with the exception of  $R_{xyz}$  which is constant at  $2.26433$ . Thus only  $R_{xyz}$  is non-zero at the origin; at  $(1, 1, 1)$   $R_{xxx} = 16.8519$ , and  $R_{xxx} = 5.61731$ . The corresponding values of  $R_{xxx}^*$  and  $R_{xxx}^*$  are  $-0.276734$  and  $-0.0922448$  respectively. At  $(1, 1, 1)$   $|R_{xxx}^*| = |Q_{xx}^*|$  and  $|R_{xxx}| = |Q_{xx}|$ , although the dimensions differ.

The symmetry of  $\text{CH}_4$  is reflected in the various electric field components and electric field gradients. At points along the C-H bond direction  $E_x = E_y = E_z$ .

Table 7. Potential, electric field and electric field gradient at points  $(a, a, a)$  for  $\text{CH}_4$

$a$	$V$	$V^*$	$E_x$	$E_x^*$	$q_{xz}$	$q_{xz}^*$
1.5	9.70321 (-1)	7.38475 (-1)	1.83964 (0)	1.41718 (0)	-6.63272 (0)	-5.73344 (0)
2.0	1.14840 (-1)	1.04007 (-1)	1.23752 (-1)	9.88716 (-2)	-2.46351 (-1)	-1.86736 (-1)
2.5	3.22414 (-2)	3.20187 (-2)	2.17703 (-2)	2.11258 (-2)	-2.89986 (-2)	-2.71059 (-2)
3.0	1.34171 (-2)	1.34153 (-2)	6.89731 (-3)	6.89095 (-3)	-6.81782 (-3)	-6.79531 (-3)
3.5	6.67148 (-3)	6.67148 (-3)	2.82885 (-3)	2.82883 (-3)	-2.28453 (-3)	-2.28453 (-3)
4.0	3.70789 (-3)	3.70789 (-3)	1.34446 (-3)	1.34446 (-3)	-9.23878 (-4)	-9.23878 (-4)
4.5	2.22998 (-3)	2.22998 (-3)	7.07972 (-4)	7.07972 (-4)	-4.24609 (-4)	-4.24609 (-4)
5.0	1.42316 (-3)	1.42316 (-3)	4.02384 (-4)	4.02384 (-4)	-2.14474 (-4)	-2.14474 (-4)
5.5	9.51476 (-4)	9.51476 (-4)	2.42691 (-4)	2.42691 (-4)	-1.16529 (-4)	-1.16529 (-4)
6.0	6.60445 (-4)	6.60445 (-4)	1.53521 (-4)	1.53521 (-4)	-6.71104 (-5)	-6.71104 (-5)
6.5	4.72856 (-4)	4.72856 (-4)	1.00997 (-4)	1.00997 (-4)	-4.05391 (-5)	-4.05391 (-5)
7.0	3.47481 (-4)	3.47481 (-4)	6.86639 (-5)	6.86639 (-5)	-2.54851 (-5)	-2.54851 (-5)
7.5	2.61087 (-4)	2.61087 (-4)	4.80070 (-5)	4.80070 (-5)	-1.65737 (-5)	-1.65737 (-5)

Table 8. Spherical Gaussian basis set for water. The molecule lies in the  $xz$ -plane with O at the origin

Exponent	$x$	$y$	$z$
9.14599 (+1)	0.00000 (0)	0.00000 (0)	0.00000 (0)
1.34143 (+1)	0.00000 (0)	0.00000 (0)	0.00000 (0)
4.64512 (-1)	7.88531 (-1)	0.00000 (0)	6.05061 (-1)
4.64512 (-1)	-7.88531 (-1)	0.00000 (0)	6.05061 (-1)
4.74836 (-1)	0.00000 (0)	0.00000 (0)	-2.52360 (-1)
1.01826 (0)	0.00000 (0)	5.00000 (-2)	0.00000 (0)
1.01826 (0)	0.00000 (0)	-5.00000 (-2)	0.00000 (0)

the diagonal elements of the electric field gradient tensor are zero and the off diagonal elements are equal. The values of the non-zero properties are listed in Table 7.

### Water

The geometry employed for the water corresponds to an internal HÔH angle of 105° and an O—H bond length of 1.81417. The molecule lies in the *xz*-plane with the oxygen nucleus at the origin and the positive *z*-axis bisects the internal HÔH angle. The basis set consists of seven spherical Gaussians and is defined in Table 8.

The density matrix and point charge model are given in Tables 9 and 10.

Table 9. Density matrix for H<sub>2</sub>O (lower triangle)

6.10934 (-2)						
2.07973 (-1)	7.14320 (-1)					
5.16081 (-4)	1.60914 (-4)	2.12016 (0)				
5.16081 (-4)	1.60914 (-4)	-1.58862 (-1)	2.12016 (0)			
6.86706 (-2)	-2.93275 (-1)	-1.57053 (0)	-1.57053 (0)	3.09698 (0)		
1.42753 (-2)	3.53045 (-2)	7.11939 (-2)	7.11939 (-2)	-1.06664 (-2)	9.84933 (+1)	
1.42735 (-2)	3.53045 (-2)	7.11939 (-2)	7.11939 (-2)	-1.06664 (-2)	-9.84193 (+1)	9.84933 (+1)

Table 10. Point charge model for water

Charge	x	y	z
3.97673 (+2)	0.00000 (0)	0.00000 (0)	0.00000 (0)
1.00000 (0)	1.43928 (0)	0.00000 (0)	1.10440 (0)
1.00000 (0)	-1.43928 (0)	0.00000 (0)	1.10440 (0)
6.98333 (-5)	3.98460 (-3)	0.00000 (0)	3.05749 (-3)
-8.91195 (-5)	2.63915 (-2)	0.00000 (0)	2.02509 (-2)
-4.24031 (0)	7.88531 (-1)	0.00000 (0)	6.05061 (-1)
6.98333 (-5)	-3.98460 (-3)	0.00000 (0)	3.05749 (-3)
-8.91195 (-5)	-2.63915 (-2)	0.00000 (0)	2.02509 (-2)
3.56622 (-1)	0.00000 (0)	0.00000 (0)	6.05061 (-1)
-4.24031 (0)	-7.88531 (-1)	0.00000 (0)	6.05061 (-1)
1.44684 (-2)	0.00000 (0)	0.00000 (0)	-1.30342 (-3)
2.49613 (-1)	0.00000 (0)	0.00000 (0)	-8.62759 (-3)
4.56760 (0)	3.89932 (-1)	0.00000 (0)	1.71639 (-1)
4.57670 (0)	-3.89932 (-1)	0.00000 (0)	1.71639 (-1)
-6.19396 (0)	0.00000 (0)	0.00000 (0)	-2.52360 (-1)
-5.43072 (-3)	0.00000 (0)	5.50543 (-4)	0.00000 (0)
-5.16374 (-2)	0.00000 (0)	3.52766 (-3)	0.00000 (0)
-1.85512 (-1)	2.47025 (-1)	3.43364 (-2)	1.89549 (-1)
-1.85512 (-1)	-2.47025 (-1)	3.43364 (-2)	1.89549 (-1)
3.75393 (-2)	0.00000 (0)	3.40990 (-2)	-8.02557 (-2)
-1.96987 (+2)	0.00000 (0)	5.00000 (-2)	0.00000 (0)
-5.43072 (-3)	0.00000 (0)	-5.50543 (-4)	0.00000 (0)
-5.16374 (-2)	0.00000 (0)	-3.52766 (-3)	0.00000 (0)
-1.85512 (-1)	2.47025 (-1)	-3.43364 (-2)	1.89549 (-1)
-1.85512 (-1)	-2.47025 (-1)	-3.43364 (-2)	1.89549 (-1)
3.75393 (-2)	0.00000 (0)	-3.40990 (-2)	-8.02557 (-2)
-1.96987 (+2)	0.00000 (0)	-5.00000 (-2)	0.00000 (0)

It is important to note that this model of water differs markedly from all previous point charge models. The point charge corresponding to each lone pair has the value  $-197$  whereas earlier models had a maximum value of  $-2$  corresponding to the two electrons. This large charge is neutralized by the large charge of  $398$  on the oxygen nucleus. The result of this feature is that the electric field around the lone pair has a much stronger angular dependence than is given by earlier models.

Recent calculations on the water dimer [4-6] have shown that the most favourable configuration for the approach of the two monomers is along a line containing an O-H bond of one and which bisects the external H $\ddot{O}$ H bond angle of the other. With these calculations in mind we have calculated one-electron properties along the line of an O-H bond and along the negative z-axis of our molecule.

The dipole moment for H<sub>2</sub>O based on the present wave function is  $0.275483$ . The low symmetry of the molecule dictates that  $Q_{xx}$  and  $Q_{yy}$  are constant,  $Q_{xy}$  and  $Q_{yz}$  are zero in the two directions investigated.  $Q_{xz}$  is also zero along the negative z-axis. The values of  $Q_{xx}$ ,  $Q_{xx}^*$ ,  $Q_{yy}$ , and  $Q_{yy}^*$  are  $-3.36956$ ,  $0.213646$ ,  $-4.56893$ , and  $-0.985721$  respectively, the error in the point charge values being  $-3.58321$ . As in the previous examples  $Q_{xz}$  in the O-H bond direction is proportional to the distance from the oxygen nucleus where it is zero. At a point  $P$  on the O-H bond, with co-ordinates  $(0.651613, 0.0, 0.5)$ ,  $Q_{xz}^*$  is  $-0.179508$  and agrees exactly with  $Q_{xz}$ .  $Q_{zz}$  is non-zero at the origin and has a value of  $-4.26952$  whilst  $Q_{zz}^* = -0.686313$ . However at  $P$  we find  $Q_{zz}$  and  $Q_{zz}^*$  have both decreased by  $0.275483$ , which is the numerical value of the dipole moment, doubling the distance from O causes a drop in both properties of twice this value. The values of  $Q_{zz}$  and  $Q_{zz}^*$  along the z-axis are tabulated in Table 11.

The third moments  $R_{xxy}$ ,  $R_{yyx}$ ,  $R_{yzz}$ , and  $R_{zyz}$  are all zero in the O-H bond direction, and, on the negative z-axis, only  $R_{yyz}$ ,  $R_{zzz}$  and  $R_{xxz}$  are non-zero.  $R_{xxx}$  and  $R_{yyy}$  have a linear dependence on the distance from the oxygen nucleus along the O-H direction. Both are zero at the origin and at  $P$  have values of  $6.58694$  and  $2.97717$  respectively.  $R_{xxx}^*$  and  $R_{yyy}^*$  are  $-0.417643$  and  $0.642309$  at the same

Table 11. Second moments for H<sub>2</sub>O along the z-axis

$z$	$Q_{zz}$	$Q_{zz}^*$
0.0	-4.26952	-0.686313
-0.8	-3.82875	-0.245541
-1.6	-3.38797	0.195231
-2.4	-2.94720	0.636003
-3.2	-2.50643	1.07678
-4.0	-2.06566	1.51755
-4.8	-1.62489	1.95832
-5.6	-1.18411	2.39909
-6.4	-0.743341	2.83986
-7.2	-0.302569	3.28063
-8.0	0.138203	3.72141

Table 12. Third moments for H<sub>2</sub>O at points in the O-H direction.  
Molecule is in the xz-plane

x	z	R <sub>zzz</sub>	R <sub>zzz</sub> <sup>*</sup>	R <sub>zzz</sub>	R <sub>zzz</sub> <sup>*</sup>	R <sub>zzz</sub>	R <sub>zzz</sub> <sup>*</sup>
0.00000 (0)	0.00000 (0)	-2.07733 (0)	1.03513 (0)	5.77349 (-1)	1.61483 (0)	0.00000 (0)	0.00000 (0)
6.51613 (-1)	5.00000 (-1)	4.53356 (0)	2.27121 (0)	2.37910 (0)	1.62498 (0)	2.96158 (0)	6.26718 (-1)
1.30323 (0)	1.00000 (0)	1.15577 (+1)	3.92051 (0)	4.41479 (0)	1.86907 (0)	6.28218 (0)	1.61245 (0)
1.43928 (0)	1.10440 (0)	1.30764 (+1)	4.31702 (0)	4.86935 (0)	1.94955 (0)	7.02080 (0)	1.86357 (0)
1.95484 (0)	1.50000 (0)	1.89950 (+1)	5.98304 (0)	6.68441 (0)	2.34709 (0)	9.96179 (0)	2.95720 (0)
2.60645 (0)	2.00000 (0)	2.68456 (+1)	8.45879 (0)	9.18798 (0)	3.05906 (0)	1.40004 (+1)	4.66097 (0)
3.25806 (0)	2.50000 (0)	3.51094 (+1)	1.13478 (+1)	1.19255 (+1)	4.00496 (0)	1.83981 (+1)	6.72375 (0)

Table 13. Potential in H<sub>2</sub>O

x	z	V	V <sup>*</sup>
Points in the O-H direction			
1.95484 (0)	1.50000 (0)	6.31140 (-1)	5.12087 (-1)
2.60645 (0)	2.00000 (0)	1.05472 (-1)	1.02299 (-1)
3.25806 (0)	2.50000 (0)	3.96774 (-2)	3.96479 (-2)
3.90968 (0)	3.00000 (0)	2.04001 (-2)	2.04000 (-2)
4.56129 (0)	3.50000 (0)	1.22863 (-2)	1.22863 (-2)
5.21290 (0)	4.00000 (0)	8.17280 (-3)	8.17280 (-3)
Points on the z-axis			
-2.4	-6.02108 (-2)	-6.78693 (-2)	
-3.2	-3.37610 (-2)	-3.38545 (-2)	
-4.0	-2.04351 (-2)	-2.04355 (-2)	
-4.8	-1.37399 (-2)	-1.37399 (-2)	
-5.6	-9.89456 (-3)	-9.89456 (-3)	
-6.4	-7.47303 (-3)	-7.47303 (-3)	

Table 14. Non-zero components of the electric field in H<sub>2</sub>O

x	z	E <sub>x</sub>	E <sub>x</sub> <sup>*</sup>	E <sub>x</sub>	E <sub>x</sub> <sup>*</sup>
Points in the O-H direction					
1.95484 (0)	1.50000	1.49387 (0)	1.12944 (0)	1.16073 (0)	8.83103 (-1)
2.60645 (0)	2.00000	1.26042 (-1)	1.13376 (-1)	1.01535 (-1)	9.18403 (-2)
3.25806 (0)	2.50000	2.90678 (-2)	2.89184 (-2)	2.39131 (-2)	2.37985 (-2)
3.90968 (0)	3.00000	1.10361 (-2)	1.10355 (-2)	8.98255 (-3)	8.98212 (-3)
4.56129 (0)	3.50000	5.29439 (-3)	5.29439 (-3)	4.17236 (-3)	4.17236 (-3)
5.21290 (0)	4.00000	2.94017 (-3)	2.94017 (-3)	2.21276 (-3)	2.21276 (-3)
Points on the z-axis					
-2.4	-2.4			3.42266 (-2)	7.13469 (-2)
-3.2	-3.2			2.39815 (-2)	2.45600 (-2)
-4.0	-4.0			1.12878 (-2)	1.12908 (-2)
-4.8	-4.8			6.15031 (-3)	6.15031 (-3)
-5.6	-5.6			3.73397 (-3)	3.73397 (-3)
-6.4	-6.4			2.44267 (-3)	2.44267 (-3)

point. At the origin  $R_{yyz}$  is  $-1.03766$  and  $R_{yyz}^*$  is  $-0.000172836$ , at  $P$  the corresponding values are  $1.24681$  and  $0.492688$ . The values of  $R_{zzx}$ ,  $R_{xxz}$  and  $R_{zzz}$  and point charge values at points in the O-H direction are presented in Table 12.  $R_{yyz}$  has a linear dependence with distance along the negative  $z$ -axis as well as in the O-H bond direction, at  $(0, 0, -0.8)$   $R_{yyz}$  is  $-4.69280$  and  $R_{yyz}^*$  is  $-0.788750$ , the non-zero values of both at the origin have been given above. In Tables 13-15 we give the values of the potential, non-zero electric fields and field gradients.

#### 4. Discussion

For each of the molecules we have examined it is possible to derive the observed relationship between the various multipole moments from the definition given in (8). The relationship between  $Q_{zz}$  at points in the O-H direction for  $H_2O$  and the dipole moment, is perhaps fortuitous, but can be derived from the appropriate expression. Since the errors produced by the point charge model in the multipole moments are closely related it should be possible to construct the quadrupole and octupole moment tensors using Buckingham's definitions [7], with the knowledge that given elements are exact and the error in others is precisely known. Naturally, knowledge of the basis set and density matrix from which the point charge model was obtained permits the exact evaluation of the multipole moments from the point charge model. The advantage is that the calculation based on the point charge model is easier to perform than that based on the original wave function.

The results presented for the potential and its derivatives show that the error involved in the point charge model is small and diminishes rapidly. Whereas the multipole moments are, in most cases, well represented by the point charge model at any point in space whether or not that point is within the confines of the nuclear framework, the potential and its derivatives are only accurate outside the shell of the molecule [1]. A contributory factor to this feature is the inability of the point charge model to include terms which arise from the Dirac delta functions which appear in  $q_{\lambda\lambda}$ .

One question that must be answered concerns the accuracy to which molecular properties can be calculated using the spherical Gaussian model. Recent work by Dixon and Tait [8] suggests that even the simplest form of spherical Gaussian model provides reliable information about molecular structure. It is known that the Frost model [9] gives about 84% of the Hartree-Fock energy, and there is evidence [8] to support the thesis that one-electron properties can be obtained as a more-or-less constant percentage of the values obtained in more accurate calculations.

#### 5. Conclusion

From the results presented in Tables 1-15 we conclude that the point charge model proposed by Hall [1] may be used to calculate one-electron properties with a high degree of accuracy. Where errors occur it is possible to evaluate them exactly without recourse to the complete spherical Gaussian wave function.

Table 15. Electric field gradients in H<sub>2</sub>O

x	z	$q_{xx}$	$q_{zz}^*$	$q_{yz}$	$q_{yy}^*$
Points in the O-H direction					
1.95484 (0)	1.50000 (0)	-3.10260 (0)	-2.52093 (0)	3.51082 (0)	2.86440 (0)
2.60645 (0)	2.00000 (0)	-1.51141 (-1)	-1.26689 (-1)	1.76179 (-1)	1.48780 (-1)
3.25806 (0)	2.50000 (0)	-2.14951 (-2)	-2.11371 (-2)	2.61621 (-2)	2.57584 (-2)
3.90968 (0)	3.00000 (0)	-5.90026 (-3)	-5.89789 (-3)	7.36033 (-3)	7.35854 (-3)
4.56129 (0)	3.50000 (0)	-2.22498 (-3)	-2.22498 (-3)	2.78651 (-3)	2.78651 (-3)
5.21290 (0)	4.00000 (0)	-1.02391 (-3)	-1.02391 (-3)	1.26434 (-3)	1.26434 (-3)
x	z	$q_{xx}$	$q_{zz}^*$	$q_{yz}$	$q_{yy}^*$
1.95484 (0)	1.50000 (0)	-4.08224 (-1)	-3.43475 (-1)	-5.08212 (0)	-4.14661 (0)
2.60645 (0)	2.00000 (0)	-2.50364 (-2)	-2.20908 (-2)	-2.55363 (-1)	-2.15685 (-1)
3.25806 (0)	2.50000 (0)	-4.66504 (-3)	-4.62126 (-3)	-3.79656 (-2)	-3.73817 (-2)
3.90968 (0)	3.00000 (0)	-1.46130 (-3)	-1.46064 (-3)	-1.06463 (-2)	-1.06431 (-2)
4.56129 (0)	3.50000 (0)	-5.61523 (-4)	-5.61525 (-4)	-3.99915 (-3)	-3.99915 (-3)
5.21290 (0)	4.00000 (0)	-2.40437 (-4)	-2.40437 (-4)	-1.79536 (-3)	-1.79536 (-3)
z	$q_{xx}$	$q_{zz}^*$	$q_{yz}$	$q_{yy}$	$q_{yy}^*$
Points on the z-axis					
-2.4	-1.07364 (-2)	-7.30841 (-2)	2.07224 (-2)	-4.08633 (-2)	1.13947 (-1)
-3.2	-1.56998 (-2)	-1.69078 (-2)	-9.34760 (-3)	-1.05507 (-2)	2.74585 (-2)
-4.0	-5.71490 (-3)	-5.72222 (-3)	-3.84941 (-3)	-3.85673 (-3)	9.57895 (-3)
-4.8	-2.45546 (-3)	-2.45548 (-3)	-1.74199 (-3)	-1.74200 (-3)	4.19749 (-3)
-5.6	-1.23112 (-3)	-1.23112 (-3)	-9.04262 (-4)	-9.04262 (-4)	2.13538 (-3)
-6.4	-6.87201 (-4)	-6.87201 (-4)	-5.17403 (-4)	-5.17403 (-4)	1.20460 (-3)

A molecular dynamics study of liquid water by Rahman and Stillinger [10] has stimulated interest in point charge models. They have used a simple point charge model for the water molecule to provide the electrostatic contribution to the effective pair potential. The point charge model examined in the present paper offers the possibility of providing a highly accurate model from which this contribution may be calculated without substantially modifying the molecular dynamics program.

Another possible application of the present point charge model is the investigation of problems of chemical reactivity. Bonaccorsi, Pullman, Scrocco, and Tomasi [11] have shown that the study of isopotential curves for large heterocycles reveals a discrimination between the different possible sites of protonation on the same molecular skeleton. They suggest that such studies may provide a method of comparing protonation sites of different kinds and also sites on different molecules. The key to such investigations is an efficient means of calculating the potential at many points around the molecule; the point charge model is well suited to this task.

*Acknowledgements.* The authors wish to thank Dr. G. W. Schnuelle for the wave function for  $H_2O$ . Financial support from SRC is also acknowledged.

### References

1. Hall, G. G.: *Chem. Phys. Letters* **6**, 501 (1973).
2. Matsuoka, O.: *Intern. J. Quantum Chem.* **5**, 1 (1971).
3. Arrighini, G. P., Guidotti, C., Maestro, M., Moccia, R., Salvetti, O.: *J. Chem. Phys.* **49**, 2224 (1968).
4. Dierksen, G. H. F.: *Chem Phys. Letters* **4**, 373 (1969).
5. Bonaccorsi, R., Petrongolo, G., Scrocco, E., Tomasi, J.: *Theoret. Chim. Acta (Berl.)* **20**, 331 (1971).
6. Dierksen, G. H. F.: *Theoret. Chim. Acta (Berl.)* **21**, 335 (1971).
7. Buckingham, A. D.: *J. Chem. Phys.* **30**, 1580 (1958).
8. Dixon, M., Tait, A. D.: To be published 1973.
9. Frost, A. A.: *J. Chem. Phys.* **17**, 3707, 3714 (1967).
10. Rahman, A., Stillinger, F. H.: *J. Chem. Phys.* **55**, 3336 (1971).
11. Bonaccorsi, R., Pullman, A., Scrocco, E., Tomasi, J.: *Theoret. Chim. Acta (Berl.)* **24**, 51 (1972).

Dr. A. D. Tait  
Department of Mathematics  
The University of Nottingham  
University Park  
Nottingham NG7 2RD, England



## Etude théorique de l'isomérisation *syn-anti* dans la Formaldoxime

Daniel Liotard, Alain Dargelos et Max Chaillet

Laboratoire de Chimie Structurale, Institut Universitaire de Recherche Scientifique, Avenue Philippon  
Pau, France

Reçu le 21 mai 1973

### *Theoretical Study of the syn-anti Isomerization of Formaldoxime*

Semi empirical CNDO and "ab initio" methods are applied to the analysis of *syn-anti* isomerization mechanism of formaldoxime. Semi empirical calculations are carried out with complete geometry optimization and lead to predicted inversion barrier equal to 37.5 kcal/mole. A bicentric partitioning of the total energy and the expression of the density matrix in hybride basis set reveal the chemical origin of the shape of the potential line.

**Key words:** Oxime nitrogen inversion Formaldoxime

### Introduction

L'existence d'isomères *syn-anti* autour de la double liaison  $C=N$  dans les oximes est expérimentalement connue de longue date [23].

Le mécanisme de cette isomérisation a suscité récemment de nombreuses études expérimentales en particulier par voies spectroscopiques [1-9]. Les interprétations théoriques proposées [7-11] considèrent en général les deux modèles limites:

a) rotation pure autour de la liaison  $C=N$  entraînant une rupture de la liaison  $\pi$

b) inversion dans le plan moléculaire via un état de transition où l'azote est hybridé *sp*.

De surcroît un changement de l'état de multiplicité de la fonction d'onde au cours du mécanisme ne peut être écarté *a priori* [4, 12].

En fait on peut définir la notion intuitive de «chemin de réaction» comme la trajectoire orthogonale (généralement unique) joignant deux cuvettes via au moins un col de l'hypersurface représentative de l'énergie totale du système étudié. Cette hypersurface est fonction d'un ensemble de coordonnées internes linéairement indépendantes qui définissent la structure moléculaire. On peut envisager autant de mécanismes que de trajectoires distinctes, leurs énergies d'activation étant respectivement fixées par «l'altitude» du col le plus haut dans le trajet suivi.

En cas de changement d'état de spin au cours de l'isomérisation un croisement des nappes (hypersurface) de multiplicité correspondante doit être observée. L'énergie d'activation peut alors ne pas être associée à un point stationnaire.

Dans de nombreux cas les trajectoires orthogonales sont correctement représentées en ne mettant en jeu qu'un nombre restreint de coordonnées internes.

Des coupes de la nappe étudiée suivant ces coordonnées présentent alors des minimums qui définissent point par point un parcours voisin passant exactement par tous les points stationnaires (cuvettes et cols).

C'est dans cet état d'esprit que nous avons jugé intéressant d'ab order le mécanisme d'isomérisation *syn-anti* de la formaldoxime en considérant la molécule isolée (phase vapeur). Tous les calculs ont été effectués par la méthode CNDO II dans son formalisme original [13]. Les résultats obtenus par voie semi-empirique ont été confrontés à un traitement «*ab initio*» comparable.

## 1. Etude préliminaire

Nous avons retenu comme topologie de la forme la plus stable de la molécule celle déterminée expérimentalement par micro onde [15] qui situe le groupement OH en position trans (Fig. 1). Ce résultat est d'ailleurs en accord avec les études conformationnelles antérieures conduites en «*ab initio*» [24, 25].

La symétrie moléculaire permet de penser que tous les complexes intermédiaires envisageables au cours de l'isomérisation correspondent approximativement au passage du groupement OH dans un plan vertical ( $\Phi = 90^\circ$ , Fig. 1). Nous avons donc estimé pour l'inversion pure ( $\Theta = 180^\circ$ ) et la rotation pure ( $\Phi = 90^\circ$ ,  $\Theta$  expérimental) la barrière d'isomérisation en admettant que la transposition ne s'accompagnait d'aucune modification sensible de structure par rapport à la forme stable.

Pour le modèle de rotation une approche de l'état triplet a été entreprise à l'aide d'un seul déterminant de Slater dans le formalisme UHF [13b]. Cette fonction d'onde ne correspond pas à un état spectroscopiquement pur et l'énergie associée ne fournit donc qu'un ordre de grandeur de celle du premier état triplet.

Afin de s'assurer de la validité de la méthode CNDO II, les mêmes calculs concernant un état singulet ont été entrepris en «*ab initio*» [16]. La base de gaussiennes choisie (Tableau 1) est construite en utilisant

- pour les hydrogènes la base 3s proposée par Huzinaga [18] contractée en 1s

- pour les atomes de la seconde période la base (7s 3p) proposée par Roos et Siegbahn [19] contractée en (3s 1p). La base contractée est déduite d'une étude des rapports des exposants proposée par David [17].

La détermination des barrières de rotation s'avérant relativement indépendante de la base [20] nous n'avons pas tenté d'améliorer cette dernière (optimisation d'exposants, orbitales de polarisation).

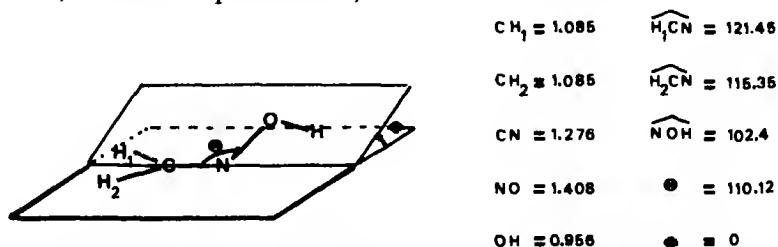


Fig. 1. Topologie moléculaire expérimentale et définition des coordonnées internes (Å et degrés décimaux)

Tableau 1a. Orbitales gaussiennes non contractées

Centre	type	Exp. Orb.	type	Exp. Orb.	type	Exp. Orb.			
C	$\gamma_1$	s	0,163484	$\gamma_5$	s	45,8498	$\gamma'_1$	x, y, z	0,199206
	$\gamma_2$	s	0,524194	$\gamma_6$	s	206,885	$\gamma'_2$	x, y, z	0,851563
	$\gamma_3$	s	3,72337	$\gamma_7$	s	1412,29	$\gamma'_3$	x, y, z	4,18286
	$\gamma_4$	s	12,3887						
N	$\gamma_1$	s	0,234424	$\gamma_5$	s	66,4630	$\gamma'_1$	x, y, z	0,286752
	$\gamma_2$	s	0,764993	$\gamma_6$	s	301,689	$\gamma'_2$	x, y, z	1,23293
	$\gamma_3$	s	5,30452	$\gamma_7$	s	2038,41	$\gamma'_3$	x, y, z	5,95461
	$\gamma_4$	s	17,8081						
O	$\gamma_1$	s	0,322679	$\gamma_5$	s	91,9805	$\gamma'_1$	x, y, z	0,365030
	$\gamma_2$	s	1,06314	$\gamma_6$	s	415,725	$\gamma'_2$	x, y, z	1,62336
	$\gamma_3$	s	7,22296	$\gamma_7$	s	2714,89	$\gamma'_3$	x, y, z	7,75579
	$\gamma_4$	s	24,4515						
H	$\gamma_1$	s	0,151374	$\gamma_2$	s	0,681277	$\gamma_3$	s	4,50038

Tableau 1b. Coefficients de contraction

Centre	type	base contractée (normalisée)					
C	1s	0,85808497	$\gamma_5$	0,18548548	$\gamma_6$	0,023955285	$\gamma_7$
	1s'	0,52909733	$\gamma_3$	0,53266497	$\gamma_4$		
	2s	0,56300794	$\gamma_1$	0,49499660	$\gamma_2$		
	2p <sub>x,y,z</sub>	0,62656913	$\gamma'_1$	0,46622710	$\gamma'_2$	0,11219402	$\gamma'_3$
N	1s	0,86171070	$\gamma_5$	0,18140919	$\gamma_6$	0,023496480	$\gamma_7$
	1s'	0,53997291	$\gamma_3$	0,52260895	$\gamma_4$		
	2s	0,57335781	$\gamma_1$	0,48615687	$\gamma_2$		
	2p <sub>x,y,z</sub>	0,61114231	$\gamma'_1$	0,47462924	$\gamma'_2$	0,11966406	$\gamma'_3$
O	1s	0,86374721	$\gamma_4$	0,17817789	$\gamma_6$	0,023878543	$\gamma_7$
	1s'	0,55097411	$\gamma_3$	0,51231799	$\gamma_4$		
	2s	0,58304895	$\gamma_1$	0,47713467	$\gamma_2$		
	2p <sub>x,y,z</sub>	0,60448414	$\gamma'_1$	0,48126911	$\gamma'_2$	0,12937303	$\gamma'_3$
H	1s	0,64766943	$\gamma_1$	0,40788964	$\gamma_2$	0,070479938	$\gamma_3$

Les différents résultats sont donnés dans le Tableau 2. Le modèle d'inversion pure apparaît nettement favorisé devant la rotation hors du plan, en dépit de l'abaissement résultant d'un changement de multiplicité de l'état de transition. Les valeurs obtenues s'inscrivent dans la ligne des résultats antérieurs [8]; l'accord entre méthode «*ab initio*» et CNDO II est une confirmation de l'aptitude de cette méthode semi-empirique à aborder les problèmes d'isomérisation.

## 2. Optimisation topologique

En premier lieu nous avons recherché, à l'aide de la méthode de plus grande pente [14], les deux cuvettes de potentiel respectivement associées aux formes cis et trans de la formaldoxime. Les deux structures stables s'avèrent planes, la forme trans étant favorisée (lignes 1 et 10 du Tableau 3). Le passage de la forme

Tableau 2. Charges nettes et moment dipolaire pour la forme Trans Energies moléculaires et barrières correspondantes

	« <i>ab initio</i> »	CNDO (geo exp)	CNDO (geo opt)
Charges nettes (en base orthogonale)			
C	0,170	0,040	0,006
N	-0,091	-0,016	0,010
O	-0,108	-0,200	-0,156
H	0,132	0,139	0,125
H <sub>1</sub>	-0,055	0,023	0,007
H <sub>2</sub>	-0,048	0,016	0,009
Moment dipolaire (D)			
Charges seules			
$\Gamma_x$	0,5003	-0,0846	0,3534
$\Gamma_y$	0,0258	-0,4477	-0,2682
$\Gamma_z$	0,5009	0,4556	0,4436
Charges et polarisations atomiques			
$\Gamma_x$	0,4992	-0,2319	0,2125
$\Gamma_y$	0,3263	-0,2330	-0,094
$\Gamma_z$	0,5964	0,3288	0,2127
(Expérimental 0,44 D [16])			
Energies (UA)			
Forme plane trans	-168,330204	-39,333185	-39,39319
Interm. Rotation (singulet)	-168,163062	-39,168631	—
(triplet)		-39,191535	-39,27277
Interm. Inversion (singulet)	-168,251960	-39,252570	-39,33343
Barrières (kcal/mole)			
Rotation (singulet)	103,2	104,8	
Rotation (triplet)		82,8	75,5
Inversion (singulet)	50,6	49,1	37,5

Tableau 3. Coupes suivant les cones d'angle au sommet  $\Theta$  (notations suivant la Fig. 1; précision:  $\pm 0,002$  Å,  $\pm 0,5$  )

$\Theta$ (degrés)	C-H <sub>1</sub> (Å)	C-H <sub>2</sub> (Å)	$\overline{\text{N-C-H}_1}$ (degrés)	$\overline{\text{N-C-H}_2}$ (degrés)	C=N (Å)	N-O (Å)	O-H (Å)	$\overline{\text{N-O-H}}$ (degrés)	E. totale (U.A.)
115,3	1,112	1,109	123,6	120,3	1,286	1,275	1,034	107,4	-39,39319
120	1,112	1,109	123,6	120,3	1,286	1,273	1,034	107,4	-39,39237
140	1,112	1,109	123,6	120,3	1,278	1,266	1,034	107,4	-39,37404
160	1,113	1,110	123,6	120,3	1,267	1,258	1,036	107,4	-39,34551
170	1,114	1,113	123,6	120,3	1,266	1,257	1,038	107,4	-39,33567
180	1,115	1,115	123,6	123,6	1,263	1,256	1,041	108,6	-39,33343
190	1,112	1,113	120,3	124,0	1,264	1,257	1,038	111,7	-39,33986
200	1,108	1,112	120,3	124,3	1,269	1,258	1,037	112,5	-39,35232
220	1,108	1,111	120,3	124,3	1,279	1,264	1,037	111,1	-39,38031
241,8	1,114	1,116	120,3	124,3	1,289	1,275	1,037	110,6	-39,39282

trans à la forme cis s'accompagne d'un léger allongement de la liaison C=N et d'une sensible réorganisation du groupement CH<sub>2</sub> (allongement des liaisons C-H, glissement du groupement dans le plan) sous l'influence répulsive du proton du groupe OH.

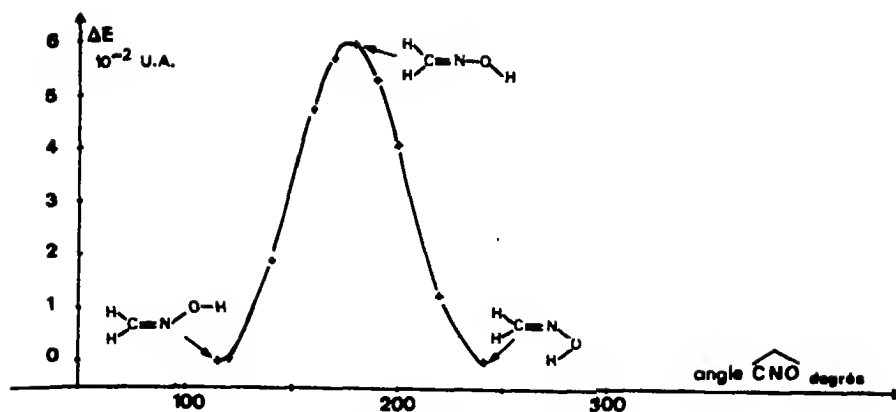


Fig. 2. Barrière d'inversion

Un mécanisme d'inversion est raisonnablement décrit à l'aide des seuls angles  $\Theta$  et  $\Phi$  (Fig. 1). Nous avons donc entrepris deux familles de coupes sur la nappe singulet:

a) angle  $\Phi$  imposé entre  $0^\circ$  et  $180^\circ$  afin de mettre en évidence un éventuel mécanisme faisant intervenir la torsion.

b) angle  $\Theta$  imposé entre  $115,3^\circ$  et  $241,8^\circ$  (valeurs correspondant aux deux formes stables de la molécule) afin de préciser un mécanisme où prédominerait l'inversion.

Les coupes suivant l'angle  $\Phi$ , dont les résultats ne sont pas reportés ici<sup>1</sup>, définissent une trajectoire symétrique dont la branche ascendante peut être décrite comme suit:

1)  $0^\circ < \Phi < 45^\circ$  Torsion pratiquement pure accompagnée d'une très faible réorganisation du reste de la molécule.

2)  $45^\circ < \Phi < 60^\circ$  Accroissement rapide de l'angle  $\Theta$  vers la valeur  $180^\circ$ .

3)  $60^\circ < \Phi < 90^\circ$  Palier horizontal correspondant à une structure plane unique où les atomes CNO sont alignés ( $\Theta = 180^\circ$ ).

Un tel type de trajectoire met clairement en évidence l'absence de complexe intermédiaire (col) de structure non plane.

Les coupes suivant l'angle  $\Theta$  (Tableau 3) confirment ce résultat. La molécule reste plane le long du chemin de réaction (Fig. 2) et le col unique obtenu correspond à la même structure moléculaire que celle du palier horizontal précédent.

La barrière d'inversion (37,5 kcal/mole) est sensiblement plus basse que celle calculée à l'aide des modèles préliminaires. Ceci semble plus raisonnable compte tenu des données expérimentales disponibles sur des composés voisins [3, 7].

Un mécanisme de rotation ne semble donc envisageable que dans l'hypothèse d'un changement de multiplicité de la fonction d'onde. A cet effet, une optimisation topologique de l'état triplet UHF a été entreprise pour la coupe  $\Phi = 90^\circ$ . La structure obtenue est d'énergie très supérieure (75,5 kcal/mole) à l'état stable

<sup>1</sup> En particulier, l'écueil d'éventuels minimums multiples suivant une coupe a été écarté en recherchant les points stationnaires à partir de plusieurs géométries initiales.

singulet (Tableau 2). Comme cette forme ne correspond pas nécessairement au point le plus haut de ce mécanisme et que la composante pure de spin déductible de la fonction d'onde serait d'énergie encore supérieure à la valeur obtenue ici, on peut considérer tout mécanisme de rotation comme thermiquement improbable.

### 3. Analyse du Mécanisme d'inversion

Afin de rechercher *a priori* l'origine chimique de la barrière d'inversion il nous a paru, devant la variété des indices statiques disponibles, préférable de retenir des quantités qui aient les dimensions d'une énergie. Pour cela, nous avons utilisé la partition bicentrique de l'énergie totale en CNDO II initialement proposée par Pople [13c] et reprise par d'autres auteurs [21]:

$$E_{\text{Tot}} = \sum_A^{\text{atome}} E_A + \sum_{B>A}^{\text{liaisons}} E_{A-B} + \sum_{B>A}^{\text{interactions}} E_{A-B} = \sum E_i.$$

L'étude de la variation de ces termes au cours de l'isomérisation peut être clarifiée par l'emploi des quantités suivantes:

coefficient de corrélation ( $E_{\text{Tot}}$ ,  $E_i$ )

- pente de la droite de corrélation correspondante.

En effet, un coefficient de corrélation voisin de l'unité indique une forte dépendance entre l'énergie  $E_{\text{Tot}}$  de la barrière et le terme  $E_i$  considéré, le poids de ce terme étant donné par la pente de la droite.

Les résultats obtenus à l'aide des coupes précédentes (Tableau 3) ont été consignés dans le Tableau 4 et appellent les remarques suivantes:

Tous les termes varient quasiment linéairement avec la barrière d'inversion, à l'exception de ceux faisant intervenir l'atome d'hydrogène du groupement OH

Tableau 4. Pente et coefficients de corrélation

Terme énergétique	Coefficient de corrélation	Pente
<b>Atomes</b>		
C	0,9758	0,207
N	0,9867	2,425
O	0,9679	0,445
H	-0,3970	-0,008
<b>Liaisons</b>		
C=N	-0,9893	-1,336
N-O	-0,9806	-1,613
O-H	0,8012	0,292
<b>Interactions</b>		
C-O	0,8874	0,194
C-H	-0,9748	-0,097
N-H	0,1390	0,118
O-(H <sub>1</sub> + H <sub>2</sub> )	0,2041	0,016
<b>Groupements</b>		
CH <sub>2</sub>	0,8971	0,309
OH	0,9806	0,728

et de quelques termes d'interaction à longue distance dont la variation moyenne est du même ordre de grandeur que la précision des calculs.

– Le terme prépondérant est nettement celui associé à l'atome d'azote, cependant que les contributions des liaisons NO et CN varient en sens contraire.

– L'ensemble des autres termes (liaisons et interactions entre voisins non liés) est identifiable à deux groupements  $\text{CH}_2$  et OH qui déstabilisent la molécule.

– Les termes d'interaction à longue distance sont tous d'importance secondaire.

Ces remarques mettent en évidence que l'origine de toute la réorganisation moléculaire au cours de l'inversion est provoquée par le changement de l'orientation des hybrides de l'atome d'azote.

En considérant que l'électronégativité d'une hybride varie comme son pourcentage d'orbitale  $s$  [22], le passage de l'état stable au complexe de transition s'accompagne

– pour la paire libre ( $sp_2 \rightarrow p$ ) d'une diminution d'électronégativité

– pour les hybrides engagées dans les liaisons ( $sp_2 \rightarrow sp$ ) d'un accroissement d'électronégativité.

Le nuage électronique entourant l'atome d'azote subit en effet cette influence comme l'indique la variation des termes de la matrice densité exprimée dans une base hybride (Tableau 5): accroissement des charges dans la direction de liaison, diminution de charge de la paire libre, légère diminution de la charge  $\pi$ .

La variation d'énergie de l'atome résulte de la somme de ces effets contraires et de la répulsion électronique qui en résulte.

Le passage ( $sp_2 \rightarrow sp$ ) accroît simultanément le caractère ionique de la liaison  $\text{C}=\text{N}$ , ce que traduit la diminution de l'indice de liaison  $\sigma$ . Cet effet est moins sensible sur la liaison  $\text{N}-\text{O}$  par suite de l'électronégativité de l'atome d'oxygène et de l'interaction des paires libres  $\sigma$  de ces deux atomes. L'effet énergétique

Tableau 5. Charges et indices de liaisons en base hybride\* (les flèches indiquent la direction où pointent les hybrides; p.l. = paire libre; A - B = Indice de liaison)

$\theta$ (degrés)	Charges				liaisons			
	$\text{N} \rightarrow \text{C}(\sigma)$	$\text{N} \rightarrow \text{O}(\sigma)$	$\text{N}(\pi)$	Np.l.( $\sigma$ )	$\text{N} - \text{C}(\sigma)$	$\text{N} - \text{O}(\sigma)$	$\text{N} - \text{C}(\pi)$	$\text{N} - \text{O}(\pi)$
115,3	1,085	0,928	1,025	1,953	0,985	0,970	0,967	0,253
120	1,090	0,942	1,023	1,937	0,984	0,968	0,966	0,258
140	1,125	0,989	1,007	1,885	0,971	0,964	0,963	0,271
160	1,154	1,008	0,988	1,878	0,965	0,974	0,961	0,277
170	1,162	1,011	0,980	1,883	0,965	0,980	0,960	0,279
180	1,162	1,015	0,983	1,876	0,964	0,980	0,961	0,277
190	1,160	1,021	0,985	1,861	0,965	0,968	0,960	0,278
200	1,151	1,024	0,992	1,847	0,966	0,957	0,960	0,279
220	1,122	1,004	1,007	1,857	0,975	0,950	0,961	0,275
241,8	1,090	0,951	1,014	1,920	0,988	0,961	0,965	0,263

\* Les hybrides  $\sigma$  sont dirigées dans les directions de liaisons et, pour la paire libre, selon la bissectrice de l'angle CNO. Leur pourcentage d'orbitale  $s$  est déduit du respect de la condition d'orthonormalisation de la base.

résultant est une stabilisation de ces liaisons dont les longueurs diminuent (Tableau 3).

Une étude graphique de la variation des termes reliés à l'atome d'hydrogène fait apparaître deux branches curvilignes correspondant aux formes *cis* et *trans* du groupement OH. L'origine de cette différenciation entre les deux formes est à rechercher dans l'anisotropie de l'entourage électronique de l'atome d'azote, ce que confirme le parallélisme des variations de la charge de la paire libre  $\sigma$  de l'azote et des termes énergétiques NH et H. Cet effet secondaire est responsable de la légère dissymétrie de la barrière d'inversion (Fig. 2) et des médiocres coefficients de corrélation obtenus (Tableau 4).

La déstabilisation du terme associé à l'atome d'azote étant en moyenne compensée par le renforcement des liaisons qu'il établit, on peut donc considérer que la hauteur de la barrière d'inversion est imposée par la nature des groupements adjacents; cette remarque suggère la possibilité d'une partition additive des barrières d'inversion en fonction des groupements rattachés au motif  $>C=N<$ ; la contribution du groupement OH apparaît plus importante que celle du groupement  $CH_3$ , en accord avec les études antérieures qui prévoient un accroissement de la barrière d'inversion sous l'influence d'hétéro-atomes porteurs de paires libres [10].

### Conclusion

L'emploi de la méthode CNDO II associée à un procédé d'optimisation topologique permet d'aborder l'étude des chemins réactionnels de manière plus satisfaisante qu'à l'aide des modèles habituellement proposés. En effet, les cuvettes et les cols de la nappe représentative de l'énergie totale du système étudié peuvent être précisés sans qu'il soit nécessaire de déterminer point par point cette hypersurface. L'ordre de grandeur des résultats obtenus, confirmé par un traitement «*ab initio*», est en accord avec les données expérimentales disponibles.

Une étude détaillée de l'origine énergétique du mécanisme peut être conduite de manière semi-quantitative à l'aide d'une partition bicentrique de l'énergie totale, en des termes beaucoup plus précis que ne le permettrait la seule analyse de la matrice densité. Cependant, l'évolution de la distribution électronique  $\sigma$  peut être clairement mise en évidence par l'emploi d'une base d'orbitales hybrides directionnelles, dans la mesure où le système étudié est raisonnablement localisable.

*Remerciements.* Ce travail a été effectué dans le laboratoire de Chimie Structurale du Professeur Deschamps que nous tenons à remercier pour son aide efficace et ses nombreux conseils. Nous remercions aussi le laboratoire de Chimie de l'E.N.S.J.F. et plus particulièrement D.J. David sans lesquels l'étude «*ab initio*» n'aurait pu être conduite.

### Références

1. Kessler, H.: *Angew. Chem. Intern. Ed. Engl.* **9**, 219 (1970)
2. Lambert, J. B., Oliver, W. L., Roberts, J. D.: *J. Am. Chem. Soc.* **87**, 5085 (1965)
3. Lefevre, J. W., Northcott, J.: *J. Chem. Soc.* 2235 (1949)
4. Curtin, D. Y., Grubbs, E. J., McCarty, G. G.: *J. Am. Chem. Soc.* **88**, 2775 (1966)



5. Vogtle, F., Mannschreck, A., Staab, H. A.: *Annalen* **708**, 61 (1967)
6. Karabatsos, G. J., Vane, F. M., Taller, R. A., Hsi, N.: *J. Am. Chem. Soc.* **86**, 3351 (1964)
7. Lehn, J. M.: *Fortschr. Chem. Forsch.* **15**, 311 (1970)
- 8a. Baechler, R. D., Mislow, K.: *J. Am. Chem. Soc.* **93**, 773 (1971)
- 8b. Cook, J., Mislow, K.: *J. Am. Chem. Soc.* **93**, 6703 (1971)
9. Raban, E., Carlson, E.: *J. Am. Chem. Soc.* **93**, 685 (1971)
10. Raban, M.: *Chem. Commun.* 1415 (1970)
11. Kerek, F., Ostrogovich, G.: *J. Chem. Soc. (B)* 541 (1971)
12. Warmb-Gerlich, D., Vogtle, F., Mannschreck, A., Staab, H. A.: *Annalen* **708**, 36 (1967)
- 13a. Pople, J. A., Santry, D. P., Segal, G. A.: *J. Chem. Phys.* **43**, 5129 (1965)
- 13b. Pople, J. A., Segal, G. D.: *J. Chem. Phys.* **43**, 5136 (1965)
- 13c. Pople, J. A., Segal, G. D.: *J. Chem. Phys.* **44**, 3289 (1966)
14. Dargelos, A., Liotard, D., Chaillet, M.: *Tetrahedron* **28**, 5595 (1972)
15. Levine, I. M.: *J. Chem. Phys.* **38**, 2326 (1963)
16. David, D. J.: IBMOL CDC 6600 version. Note technique du Centre de calcul de l'I.N.S.J.F. Paris (1969)
17. David, D. J.: *Theoret. chim. Acta (Berl.)* **23**, 226 (1971)
18. Huzinaga, S.: *J. Chem. Phys.* **42**, 1293 (1965)
19. Roos, B., Siegbahn, P.: *Theoret. Chim. Acta (Berl.)* **17**, 209 (1970)
20. Lehn, J. M.: *Proceedings of the seminar on computational problems in quantum chemistry Strasbourg, September (1969)*
- 21a. Leibovici, C.: *J. Mol. Struct.* **10**, 333 (1971)
- 21b. Leibovici, C.: *J. Mol. Struct.* (1973) à paraître
22. Hinze, J., Jaffe, H. H.: *J. Am. Chem. Soc.* **84**, 540 (1962)
23. Sauvaitre, H.: Thèse Pau (1969)
24. Radom, L., Hchre, W. J., Pople, J. A.: *J. Am. Chem. Soc.* **93**, 289 (1971)
25. Roob, M. A., Csizmadia, I. G.: *J. Chem. Phys.* **50**, 1819 (1968)

Dr. D. Liotard  
Laboratoire de Chimie Structurale  
Institut Universitaire de Recherche Scientifique  
Avenue Philippon  
Pau, France



## Etude théorique de la conformation des complexes moléculaires de type donneur accepteur d'électrons

Roger Arnaud, Danielle Faramond-Baud et Maurice Gelus

Université Scientifique et Médicale de Grenoble, Laboratoire de Chimie Générale, Centre de Tri  
F-38041, Grenoble-Cedex

Reçu le 30 mai 1973

### *Theoretical Study of the Conformation of Donor-Acceptor Molecular Complexes*

The stabilisation energy of donor acceptor molecular complexes given by tetracyanoethylene with thiazole and phenylthiazoles have been calculated by several semiempirical approaches. The results obtained with the Extended Hückel theory, the CNDO/2 method and the PCILO method, the latter a perturbative one, are compared and this comparison shows that the PCILO method is well adapted to the study of molecular complexes. The different contributions to the stabilisation energy and the origin of the intermolecular forces have been discussed in relation to the geometry of the complex. For all the studied complexes, the calculated results obtained by the PCILO method agree very well with experiments.

**Key words:** Donor-acceptor complexes, conformations of ~

### 1. Introduction

De nombreux travaux ont été consacrés ces dernières années à l'étude des forces intermoléculaires responsables de la stabilité des complexes moléculaires de type donneur-accepteur d'électrons dans leur état fondamental. La nature et la contribution relative de ces forces restent encore très controversées [1-3]. L'une des principales difficultés semble résider dans le choix d'un modèle adéquat pour aborder l'étude théorique de ces gros systèmes, et l'évaluation de l'énergie de stabilisation peut être envisagée essentiellement selon deux voies différentes. Dans le premier cas, les molécules constituant le complexe sont considérées séparément et l'énergie de stabilisation est alors calculée en faisant la somme de différentes contributions à l'interaction intermoléculaire. Citons à ce sujet, les travaux de Lippert [4], de Mantione [5] qui ont étudié les systèmes tétracyanoéthylène-méthyl benzènes. Cependant, les différentes contributions intermoléculaires restent très difficiles à évaluer et il apparaît donc intéressant de calculer l'énergie de stabilisation, selon une deuxième voie, c'est-à-dire en considérant le complexe comme un système unique formé des deux molécules. De tels calculs sont devenus possible avec le développement de méthodes semi-empiriques, incluant tous les électrons de valence. Les résultats ainsi obtenus sont nombreux [6], mais varient suivant le nombre et la grandeur des paramètres utilisés.

On sait, d'autre part, que le thiazole ou ses dérivés donnent, avec le tétracyanoéthylène (TCNE) des complexes moléculaires, dont les énergies de stabilisation, exception faite pour le système TCNE-thiazole, ont été mesurées [7, 8].

Nous nous sommes proposés d'étudier quelques uns de ces complexes par la méthode PCILO [9-11], et dans le cas du complexe TCNE-thiazole, nous avons comparé les résultats à ceux obtenus par la méthode de Hückel étendue à tous les électrons de valence (Extended Hückel theory) [12] et par la méthode CNDO [13]. Nous avons aussi tenté de déterminer à priori quelle était la conformation de ces systèmes.

## 2. Les méthodes utilisées; le choix des paramètres et des géométries

Si on considère le complexe TCNE-thiazole comme un système unique, il est hors de question de faire un calcul du type «*ab initio*». Cependant, pour étudier la conformation des molécules, il semble nécessaire de tenir compte de tous les électrons de valence. Aussi, avons nous utilisé les méthodes «extended Hückel», CNDO/2, qui sont déjà bien connues et la méthode PCILO, plus récente.

### La méthode EH

C'est la méthode la plus simple et elle a parfois été appliquée à l'étude de la forme des molécules [14].

### La méthode CNDO/2

C'est une méthode du type S.C.F. Nous avons utilisé la version CNDO/2, dont le programme a été écrit par Dobosh [15], la base d'orbitales étant la base minimum d'orbitales de Slater. Nous avons, dans ces calculs, rencontré des problèmes de convergence, qui ont été résolus en adoptant la technique proposée par Chesnut [6c].

### La méthode PCILO

La méthode PCILO est basée sur un traitement de perturbation qui introduit des termes d'énergie de corrélation [9-11].

On construit un déterminant de Slater à partir d'un ensemble raisonnable d'orbitales liantes localisées sur les liaisons chimiques. Ces orbitales liantes  $|i\rangle$  sont obtenues en faisant une combinaison linéaire d'orbitales hybrides prises deux à deux:

$$|i\rangle = C_{i1} |i_1\rangle + C_{i2} |i_2\rangle$$

$i_1$  et  $i_2$  étant les orbitales hybrides, chacune d'entre elles appartenant à l'un des deux atomes intervenant dans la liaison. A chaque orbitale liante  $|i\rangle$  est associée une orbitale antiliante  $|i^*\rangle$ , qui lui est orthogonale

$$|i^*\rangle = -C_{i2} |i_1\rangle + C_{i1} |i_2\rangle.$$

Les orbitales liantes permettent de construire la fonction d'onde d'ordre zéro et les orbitales antiliantes, des déterminants de Slater correspondants aux configurations excitées. Celles-ci se répartissent en deux catégories selon que l'excitation correspond ou non au transfert d'un électron d'une liaison à une autre. L'énergie de l'état fondamental est obtenue par un développement de perturbation du type Rayleigh-Schrödinger.

La méthode PCILO s'avère très avantageuse, car l'utilisation de la technique de perturbation à partir d'orbitales localisées, l'approximation du recouvrement différentiel nul (ZDO) réduisent considérablement le temps de calcul. De plus, elle s'avère plus adaptée que la méthode CNDO/2 aux problèmes de conformation de molécules conjuguées [16].

### *Choix des paramètres*

Dans le cas de la méthode « Extended Hückel », le facteur de proportionnalité intervenant dans le calcul des éléments non diagonaux a été pris égal à 0,75.

Pour appliquer la méthode PCILO au calcul du thiazole, il est nécessaire d'évaluer les intégrales faisant intervenir les orbitales de l'atome de soufre. Ces intégrales ont été calculées à partir des formules proposées par Malricu [17]; les orbitales  $3d$  n'ont pas été prises en compte: il est apparu, à la suite des travaux de Schwartz [18] et Hillier [19] que les orbitales  $s$  et  $p$  forment une base suffisante dans un problème de conformation.

### *Choix des géométries moléculaires*

Il a été supposé, dans tous les calculs, que la géométrie des molécules constituant le complexe était identique à celle des molécules isolées.

Pour les TCNE, la géométrie adoptée est celle déterminée par Bekoe et Trueblood [20] à partir de mesures cristallographiques, et dans le cas du thiazole sa géométrie a été définie par ailleurs [21]. Dans le cas des phényl-thiazoles le problème est différent, car leur géométrie n'a pas été déterminée expérimentalement. Nous avons considéré que la molécule de phényl-thiazole était une molécule plane. En effet, des calculs préliminaires par la méthode PCILO visant à déterminer la conformation la plus stable des monophényl-thiazoles n'ont pas permis de mettre en évidence l'existence d'une barrière de rotation dépassant 0,6 kcal/mole, énergie qui n'est pas suffisante pour empêcher, en solution, la libre rotation du groupement phényle. Ce sujet a été abordé récemment par Trinajstić et Galasso [22, 23], mais il ne semble pas qu'il ait reçu une solution définitive.

En ce qui concerne le complexe lui-même, nous avons adopté une structure de type « sandwich », les plans des molécules de TCNE et de thiazole étant maintenus parallèles.

L'énergie de stabilisation du complexe est calculée par différence entre l'énergie totale du complexe et celle obtenue en faisant la somme des énergies des deux molécules constituant le complexe. La variation de l'énergie de stabilisation est étudiée en fonction des deux paramètres possibles: la distance interplan  $R$  et la disposition relative des molécules, qui peut se caractériser par l'angle  $\theta$ , tel qu'il est défini dans la Fig. 1.

## **3. Résultats et discussions**

### *a) Etude comparée du complexe TCNE-thiazole par les méthodes EH, CNDO/2 et PCILO*

Nous avons résumé dans le Tableau 1 la variation de l'énergie de stabilisation  $\Delta E$  en fonction de la distance interplan  $R$ .

Tableau 1

$R$ (Å)	$\Delta E$ kcal/mole <sup>-1</sup> EH	$\Delta E$ kcal/mole <sup>-1</sup> CNDO/2	$\Delta E$ kcal/mole <sup>-1</sup> PCILO
1,25	-2509,2	-35,5	
1,4		-253,4	
1,5	-1510,7	-301,1	
1,6		-304,2	191,6
1,75	-834,32	-254,1	74,0
2,0	-416,8	-154,3	1,4
2,15		-2,5	-9,2
2,25	-95,9		-10,9
2,35	-19,4		-10,5
2,5			-8,5
3,0			-2,5
3,4			-1,0

L'examen de ces résultats montre qu'aucun minimum d'énergie n'est observé en appliquant la méthode «EH». La même observation a été faite par Wold [24] qui a étudié le complexe TCNE-benzène. Ceci est probablement dû aux approximations faites dans le calcul de l'énergie, que est simplement prise égale à la somme des valeurs propres de l'hamiltonien. Il ne semble donc pas que cette méthode soit convenable pour évaluer l'énergie de stabilisation d'un complexe.

Dans le cas des méthodes CNDO/2 et PCILO, au contraire, un minimum d'énergie est obtenue pour des valeurs de  $R$  valant respectivement 1,60 Å et 2,25 Å. Les énergies correspondantes étant de -304,2 et -10,9 kcal/mole. Nous ne possédons aucune donnée expérimentale sur distance interplan et sur l'énergie de formation due complexe TCNE-thiazole. Par comparaison avec des complexes voisins, on peut donner un ordre de grandeur des valeurs:  $R$  devrait être compris entre 3,2 et 3,5 Å,  $\Delta E$  entre -6 et -9 kcal/mole.

Il apparaît donc que la méthode PCILO donne une image plus proche de la réalité que la méthode CNDO/2 et il y a lieu de discuter ce résultat.

La valeur de  $\Delta E$  obtenue par la méthode CNDO/2 est nettement trop élevée et la valeur de  $R$  trop faible; ces mêmes faits avaient été notés par Chesnut [6e] dans une étude du complexe TCNE-benzène, ainsi que dans d'autres travaux concernant la liaison hydrogène [25, 26], qui met en jeu des interactions à grande distance. Il avait déjà été remarqué [13] que la méthode CNDO donnait des énergies de liaison trop grandes et qu'elle n'était pas toujours apte à rendre compte des propriétés des systèmes conjugués [16, 27]. Pour expliquer ce résultat, on peut rappeler que la méthode CNDO est basée sur le schéma SCF, que l'on sait peu adapté dans les problèmes faisant intervenir des distances plus grandes que celles mises en jeu dans les liaisons chimiques classiques. D'autre part, la méthode a été mise au point en introduisant des approximations sur les intégrales et des paramètres, qui sont certainement moins valables dans le cas d'un complexe. En dernier lieu, les calculs du type SCF, dans l'approximation LCAO, sans interaction de configuration, ne sont vraiment significatifs que s'il est possible de définir des orbitales moléculaires pour décrire le système étudié. Or, dans un complexe de transfert de charge, les deux molécules initiales restent bien individualisées, et, dans l'état fondamental, il est douteux que l'on puisse avoir une

bonne description avec des orbitales moléculaires occupées, étendues à tout le système, si la distance expérimentale présumée est de l'ordre de 3.5 Å. Il est possible que l'ensemble des résultats obtenus par la méthode CNDO/2 soit amélioré en introduisant une interaction de configuration, mais, a priori, les temps de calcul seraient très importants, et disproportionnés par rapport au but recherché.

Les résultats obtenus par la méthode PCILO sont très différents, et beaucoup plus proches de la réalité expérimentale. Ceci n'est pas surprenant, dans la mesure où la méthode PCILO, qui repose sur un traitement de perturbation, tient compte explicitement de l'énergie de corrélation, qui est d'autant plus importante que les distances sont grandes. Avant d'examiner l'origine des différentes contributions à l'énergie de stabilisation, on peut mettre en évidence le rôle que joue l'énergie de corrélation. Dans la description d'un complexe par la méthode de CNDO/2, il apparaît que les liaisons des molécules initiales, et notamment celles du TCNE, sont très fortement polarisées. Or, plus une liaison devient polaire, plus la corrélation intralialison diminue, ce qu'on peut constater en remarquant que, pour une liaison localisée polarisée, la correction due à la diexcitation intralialison s'écrit [11]:

$$\epsilon_{ii} = \frac{K_{ii}^2}{\Delta E \left( \frac{i^* i^*}{i i} \right)}$$

avec

$$K_{ii} = C_{i1}^2 C_{i2}^2 (g_{1111} + g_{1212} - 2g_{1112})$$

intégrale qui s'annule lorsque la liaison  $i$  a sa polarité égale à 1. Il est donc possible que les calculs variationnels du type SCF ne tiennent pas compte de cette diminution de la corrélation intramoléculaire au cours de la formation du complexe.

Bien que la méthode PCILO donne des résultats plus satisfaisants, elle conduit encore à une énergie de stabilisation trop grande, et à une distance interplan  $R$  trop petite. Ce résultat est probablement dû à la sous-estimation de la répulsion électronique, comme dans la méthode CNDO/2. Rien d'étonnant à cela, puisque les approximations concernant le calcul des intégrales introduites la méthode CNDO/2 sont reprises dans la méthode PCILO: le recouvrement différentiel est négligé, et toutes les intégrales atomiques sont évaluées à partir d'orbitales atomiques  $s$ , ces deux approximations concourant à sous estimer la répulsion.

Un autre point est la variation de l'énergie de stabilisation avec l'angle  $\Theta$  (Fig. 1). Là, les deux méthodes donnent des résultats identiques: pour une distance  $R$  égale à 3,4 Å, de l'ordre de la distance expérimentale, la conformation la plus stable correspond à un angle  $\Theta$  de 90°, alors que pour la distance  $R$  minimum calculée, l'angle  $\Theta$  est de 0°. Il n'est pas possible de tirer une conclusion de ces résultats, d'autant plus que des calculs effectués sur le même système, mais par une technique différente [8], donnait, pour une distance  $R$  égale à 3,4 Å, une énergie minimum pour  $\Theta = 0^\circ$ . Quoiqu'il en soit les différences d'énergies mises en jeu sont trop petites pour être significatives.

Il peut être intéressant de comparer les énergies apparaissant aux différents ordres du développement de perturbation dans la méthode PCILO et l'énergie calculée par la méthode CNDO/2, les résultats correspondants étant rassemblés dans le Tableau 2.

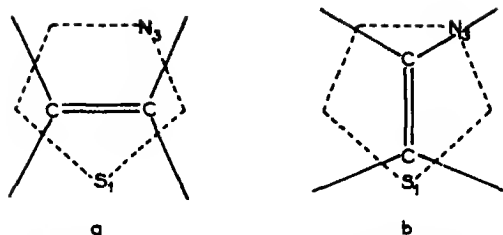
Fig. 1. Définition de l'angle  $\theta$ . a  $\theta = 0$  ; b  $\theta = 90$ 

Tableau 2

Energie* (kcal mole <sup>-1</sup> )	Thiazole	TCNE	Complexe ( $R = 2.0 \text{ \AA}$ )
Fondamental $E_0$	+ 177,01	+ 352,62	+ 779,58
$E_0$ + monoexcitations	+ 14,94	- 6,01	- 27,15
Après 2ème ordre	- 131,33	- 429,15	- 550,80
Après 3ème ordre	90,08	- 94,52	- 59,39
$E(\text{CNDO})/2$	- 28868,48	- 55335,31	- 84327,57

\* Les énergies sont exprimées par rapport à l'énergie totale obtenue par la méthode CNDO/2.

L'examen des résultats permet de faire les remarques suivantes:

Pour les molécules isolées, comme pour le complexe, la somme de l'énergie du fondamental et des contributions des monoexcitations apparaissant au deuxième ordre, est très proche de l'énergie calculée dans le schéma CNDO/2.

L'énergie au second ordre est fortement négative, l'énergie au troisième ordre est légèrement positive, et tempère l'abaissement excessif de l'énergie correspondant au second ordre. Ces remarques semblent être caractéristiques de la méthode [11] dans laquelle les contributions énergétiques ont, au moins jusqu'au deuxième ordre du traitement de perturbation, une signification physique que nous allons discuter maintenant.

#### b) Examen des différentes contributions énergétiques dans le complexe

Il nous est apparu intéressant de regarder parmi les contributions énergétiques apparaissant dans le développement de perturbation, celles qui sont «sensibles» à la variation de la distance intermoléculaire et celles qui participent à la stabilisation du complexe. Les résultats détaillés sont donnés dans le Tableau 3 pour le système TCNE-thiazole. Les remarques sont les mêmes que celles déduites de l'étude du TCNE-benzène [28] et peuvent être étendues à tous les complexes envisagés.

Pour faire apparaître à la fois la variation avec  $R$  et la participation «effective» à la stabilisation du complexe, nous donnerons pour chacune des contributions les valeurs de  $\Delta E$ ; nous faisons figurer dans la dernière colonne du tableau les énergies  $E_\infty$  (somme des contributions des deux molécules isolées) correspondantes. On peut raisonnablement penser que l'analyse ainsi faite des différents



Tableau 3

Contributions énergétiques (kcal mole <sup>-1</sup> )	R (Å)	1,75	2,0	2,15	2,25	2,35	2,50	3,0	3,4	E <sub>0</sub> (kcal mole <sup>-1</sup> )
Fondamental ΔE <sub>0</sub>		+ 310,16	+ 126,17	+ 71,95	+ 49,06	+ 32,23	+ 14,32	+ 2,30	+ 0,40	- 83674,16
m <sub>1</sub> <sup>a</sup>		0	0	0	0	0	0	0	0	- 53,50
m <sub>2</sub> <sup>b</sup>		- 418,43	- 231,72	- 155,74	- 117,63	- 87,82	- 55,54	- 10,41	- 2,39	- 467,21
d <sub>1</sub> <sup>c</sup>		0	0	0	0	0	0	0	0	- 167,90
d <sub>2</sub> <sup>d</sup>		- 14,20	- 8,53	- 6,35	- 5,25	- 4,36	- 3,33	- 1,48	- 0,84	- 401,52
Energie après le 2 <sup>e</sup> ordre ΔE <sub>2</sub>		- 122,47	- 114,08	- 90,14	- 74,92	- 59,95	- 44,55	- 9,59	- 2,83	- 84764,29
m <sub>1</sub> - d <sub>1</sub> <sup>e</sup>		0	0	0	0	0	0	0	0	+ 7,99
m <sub>1</sub> - m <sub>1</sub> <sup>e</sup>		- 0,50	- 0,42	- 0,38	- 0,36	- 0,34	- 0,31	- 0,28	- 0,23	+ 25,92
m <sub>1</sub> - m <sub>2</sub> <sup>e</sup>		- 25,61	- 15,52	- 10,96	- 8,54	- 6,57	- 4,33	- 0,91	- 0,22	+ 6,33
m <sub>1</sub> - d <sub>2</sub> <sup>e</sup>		- 0,54	- 0,46	- 0,42	- 0,39	- 0,37	- 0,33	- 0,25	- 0,19	+ 29,89
m <sub>2</sub> - d <sub>2</sub> <sup>e</sup>		- 1,53	- 0,63	- 0,35	- 0,23	- 0,15	- 0,07	- 0,01	0	+ 9,78
d <sub>1</sub> - d <sub>2</sub> <sup>e</sup>		+ 4,72	+ 2,82	+ 2,09	+ 1,73	+ 1,43	+ 1,09	+ 0,48	+ 0,27	+ 150,77
m <sub>2</sub> - m <sub>2</sub> <sup>e</sup>		+ 240,41	+ 120,89	+ 84,60	+ 65,61	+ 50,27	+ 32,99	+ 6,87	+ 1,69	- 24,99
d <sub>2</sub> - d <sub>2</sub> <sup>e</sup>		+ 15,5	+ 8,85	+ 6,37	+ 5,13	+ 4,16	+ 3,06	+ 1,2	+ 0,63	+ 189,74
Energie après le 3 <sup>e</sup> ordre ΔE		+ 73,96	+ 1,44	- 9,21	- 10,89	- 10,53	- 8,48	- 1,98	- 0,84	- 84388,40

<sup>a</sup> Energie de polarisation (monoexcitation dans une même liaison  $\left(\frac{1^* 1^*}{1}\right)$ ).

<sup>b</sup> Energie de délocalisation ou de transfert de charge (monoexcitation dans deux liaisons différentes  $\left(\frac{1^* 1^*}{1}\right)$ ).

<sup>c</sup> Energie de corrélation intra-liaison (d'excitation dans une liaison  $\left(\frac{1^* 1^*}{11^*}\right)$ ).

<sup>d</sup> Energie de corrélation inter-liaison ou de dispersion (monoexcitation dans deux liaisons différentes  $\left(\frac{1^* 1^*}{1j}\right)$ ).

<sup>e</sup> Interaction entre les différents termes mentionnés ci-dessus.

termes énergétiques nous permet de traduire en gros les interactions intermoléculaires compte tenu des hypothèses de départ: recouvrement différentiel nul et orbitales localisées.

L'analyse du Tableau 3 montre que les contributions fortement dépendantes de la distance  $R$  sont également celles dont le « poids » dans l'énergie de formation du système est important. Nous consacrerons l'essentiel de la discussion à l'examen des termes énergétiques intervenant à l'ordre 2 car il est possible d'en donner une interprétation physique

### 1. Energie de polarisation et de corrélation intra-liaison

Ces énergies sont constantes quelque soit  $R$  et ne participent pas à la stabilisation du complexe, ce qui semblerait logique puisque les géométries intramoléculaires sont figées. Paradoxalement, la corrélation intralialison est fortement sous-estimée, compte tenu des approximations faites, comme cela a été vu plus haut.

### 2. Energie de délocalisation ou de transfert de charge

Elle résulte de l'excitation du type  $i \rightarrow j^*$  qui conduit à la délocalisation des électrons de la liaison  $i$  vers la région de la liaison  $j$ . Cette énergie dépend fortement de  $R$  et sa contribution à la formation du complexe est importante.

### 3. Energie de corrélation interliaison ou de dispersion

De l'analyse des résultats, il ressort que ce terme varie peu en fonction de la distance intermoléculaire  $R$ . Effectivement, compte tenu de l'approximation ZDO, seules les contributions de diexcitations sans transfert de charge représentées par les intégrales  $(ij|i^*j^*)$  interviennent, les intégrales  $(ij|k^*l^*)$  correspondant à des interactions avec transfert de charge étant nulles. Une diexcitation interliaison correspond alors à une double monoexcitation  $i \rightarrow i^*$  et  $j \rightarrow j^*$ . L'interaction entre les distributions dipolaires  $ii^*$  et  $jj^*$  diminue comme  $r^{-3}$  quand la distance  $r$  entre les liaisons  $i$  et  $j$  augmente. L'expression de l'énergie de corrélation inter-liaison, est donnée par [11]:

$$d_2 = \sum_i \sum_{j < i} \frac{4(ij|i^*j^*)^2}{E_0 - E\left(\begin{smallmatrix} i^* & j^* \\ ij \end{smallmatrix}\right)}$$

avec

$$E_0 - E\left(\begin{smallmatrix} i^* & j^* \\ ij \end{smallmatrix}\right) = e_{j^*} + e_{i^*} - e_j - e_i - J_{ii^*} - J_{jj^*} + K_{ii^*} + K_{jj^*}.$$

Le terme du dénominateur correspond donc à la somme  $\Delta_{jj^*}$  et  $\Delta_{ii^*}$  énergies de polarisation indépendantes des variations de la géométrie intermoléculaire. Quant au numérateur, il comprend les intégrales coulombiennes calculées dans le cadre CNDO/2 c'est-à-dire sur des orbitales  $s$  de Slater: ces intégrales ne varient notablement avec  $R$  que pour des distances interatomiques classiques, leur variation pour des distances supérieures à 2 Å est très faible. Ceci explique qu'à partir de 2,5 Å,  $d_2$  décroît beaucoup moins vite que  $m_2$  quand  $R$  augmente. La contribution inter-liaisons est donc faible; or il s'agit dans notre cas d'une interaction dipôle-dipôle où  $i$  et  $j$  sont définis sur deux liaisons différentes et de plus très éloignées.

Nous remarquons enfin que la correction la plus importante, au troisième ordre, est fournie par l'énergie de délocalisation-délocalisation et que cette correction est positive.

*c) Variation de l'énergie de stabilisation pour les systèmes TCNE-phényl thiazoles*

Les résultats obtenus par la méthode PCILO pour le complexe TCNE-thiazole étant convenables, nous avons envisagé l'étude des complexes du TCNE avec quelques dérivés phényles du thiazole; il était en effet intéressant de connaître «l'aptitude» de cette méthode à rendre compte de la variation de l'énergie de stabilisation dans une famille de composés. Les grandeurs expérimentales qui permettent de chiffrer la stabilité d'un complexe moléculaire  $\pi$  sont la constante de stabilité  $K$  et mieux l'enthalpie de formation  $\Delta H$ ; ces grandeurs ont été mesurées dans le cadre d'une étude en solution des systèmes TCNE-phényl-thiazoles [7]. Si l'on connaît la variation de l'énergie de formation dans les systèmes envisagés, il sera par contre plus difficile de comparer les ordres de grandeur des énergies mesurées et calculées, aucune mesure n'ayant été faite en phase gazeuse.

Pour la conformation correspondant à  $\theta = 0^\circ$ , nous avons illustré sur la Fig. 2 la variation d'énergie  $\Delta E$  en fonction de la distance intermoléculaire. Dans le Tableau 4 les résultats obtenus pour la distance d'équilibre sont résumés et

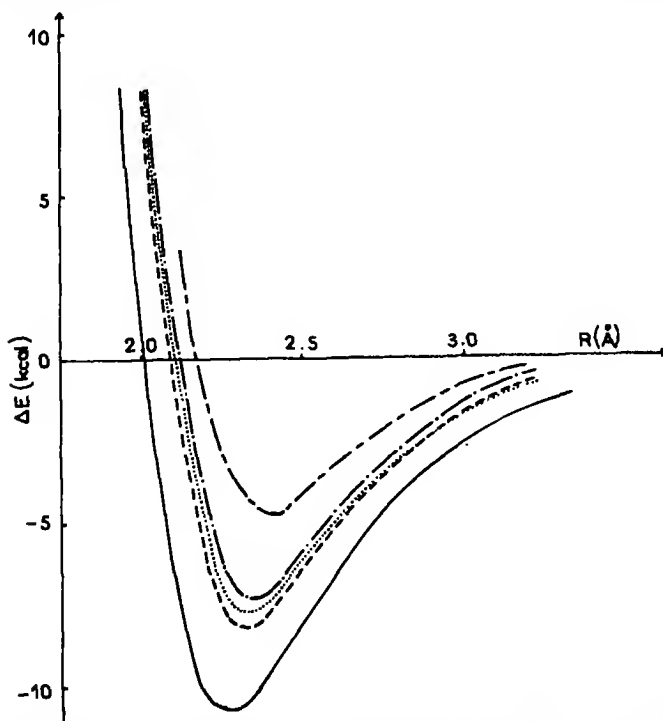


Fig. 2. Courbe d'énergie potentielle des complexes fermés entre le TCNE et différents thiazoles:  
 — thiazole — — — Phényl-4 thiazole ..... Phényl-2 thiazole - · - · - Phényl-5 thiazole · · · Di-phényl-2,4 thiazole

Tableau 4

Donneurs	Phényl-4 thiazole	Phényl-2 thiazole	Phényl-5 thiazole	Diphényl-2,4 thiazole
$R_{eq}$ (Å)	2,32	2,34	2,35	2,41
$\Delta E$ (kcal mole <sup>-1</sup> )	- 8,31	- 7,76	- 7,42	- 4,83
$K$ (l · mole <sup>-1</sup> )	6,65	4,48	4,98	3,62
$\Delta H$ (kcal mole <sup>-1</sup> )	- 3,9	- 2,9	- 3,1	- 2,5

figurent également dans ce tableau les constantes  $K$  et les enthalpies de formation  $\Delta H$  mesurées dans le tétrachlorure de carbone à 20° C.

Les énergies de stabilisation obtenues rendent assez bien compte de la variation de la «force» des complexes: compte tenu des résultats obtenus avec le thiazole, les  $\Delta E$  calculés diminuent lorsque le nombre de substituants phényle sur le cycle thiazolique augmente, ce qui est en accord avec les mesures expérimentales.

Les valeurs des énergies calculées sont voisines de celles obtenues par Lippert [4] et Mantione [3] pour une série de complexes formés entre le TCNE et des hétérocycles pentagonaux. On peut également avoir une idée approchée de l'enthalpie de formation en phase gazeuse en se référant aux travaux de Kroll [29] sur les complexes TCNE-méthylbenzènes; en l'absence d'interaction avec le solvant, l'énergie de stabilisation est fortement augmentée, à titre d'exemple, Briegleb [30] dans une étude sur le système TCNE-durène dans le tétrachlorure de carbone a obtenu les valeurs suivantes:  $K = 14,061 \cdot \text{mole}^{-1}$ ,  $\Delta H = 5,48 \text{ kcal mole}^{-1}$ , en phase gazeuse ces valeurs sont respectivement de  $118001 \cdot \text{mole}^{-1}$  et  $-10,1 \text{ kcal mole}^{-1}$ . Les valeurs des énergies calculées, quoique certainement un peu fortes, se comparent favorablement aux valeurs attendues.

#### Remarque

Les calculs ont été effectués sans itération sur les polarités; on pouvait se demander si cette simplification était légitime pour les systèmes envisagés qui ont des charges nettes assez importantes (le TCNE en particulier). Un calcul avec itération sur les polarités a été effectué sur le système TCNE-phényl-5 thiazole; pour le TCNE un nombre de cycle d'itération supérieur à 20 était nécessaire pour atteindre la convergence avec le programme original; nous l'avons accélérée en utilisant au  $i^{\text{ème}}$  cycle pour les coefficients des orbitales hybrides atomiques la demi-somme des valeurs obtenues aux cycles  $(i-2)$  et  $(i-1)$ . Les résultats avec ou sans itération sont sensiblement les mêmes:  $\Delta E_{\text{iter}} = -7,93 \text{ kcal mole}^{-1}$  (au lieu de  $-7,42 \text{ kcal mole}^{-1}$ ); nous constatons une diminution de l'énergie de polarisation mais le «poids» relatif des différentes contributions énergétiques à l'énergie de stabilisation est conservé, de ce fait le processus itératif n'a pas été retenu pour les autres calculs.

#### 4. Conclusion

Notre but n'était pas d'obtenir des résultats quantitatifs, mais de voir dans quelle mesure les méthodes actuelles permettent de rendre compte de la géométrie d'un complexe.

L'analyse des ordres relatifs, comparés aux valeurs estimées montrent que la méthode PCILO semble mieux appropriée à l'étude du problème d'interactions à moyenne distance et se révèle par sa remarquable rapidité un précieux outil de travail dans les problèmes de conformation.

Le problème intéressant était aussi de rendre compte de la nature des forces mises en jeu dans de tels complexes; la méthode CNDO/2 donne un résultat énergétique global, difficile à relier à des notions structurales. La méthode PCILO, au contraire permet d'interpréter physiquement les différentes contributions énergétiques jusqu'au 2<sup>ème</sup> ordre. Mais, du fait qu'elle est limitée par des approximations, elle ne rend pas compte des termes de polarisation, de dispersion inter-moléculaires, mais seulement de la contribution de transfert de charge.

Notons que quelle que soit la méthode de calcul variationnelle (CNDO/2 ou diverses approches SCF) ou perturbative (PCILO), les résultats conduisent toujours à une énergie de stabilisation trop élevée. Aussi se pose le problème suivant important: cet abaissement d'énergie  $\Delta E$  calculé théoriquement représente-il uniquement l'énergie d'interaction correspondant à l'expérience ( $\Delta E = E_{AD} - (E_A + E_D)$ ), ou tient-il compte aussi d'un abaissement d'énergie dû au fait que chaque molécule composante A et D dans le système moléculaire A-D profite de la présence de sa voisine pour améliorer sa description propre.

*Remerciements.* Les auteurs remercient Messieurs G. Berthier et J. P. Malrieu pour les fructueuses discussions concernant ce travail.

### Références

1. Dewar, M. J. S., Thompson, C. C. Jr.: *Tetrahedron Suppl.* **7**, 97 (1966)
2. Claverie, P., Malrieu, J. P.: *J. Chim. Phys.* **65**, 735 (1968)
3. Mantione, M. J.: Dans: Pullman, B. (Ed.): *Molecular association in biology*, p. 411. New York: Academic Press 1968
4. Lippert, J. L., Hanna, M. W., Trotter, P. J.: *J. Am. Chem. Soc.* **91**, 4035 (1969)
5. Mantione, M. J.: *Theoret. chim. Acta (Berl.)* **15**, 141 (1969)
- 6a. Sundaram, K., Prucell, W. P.: *Int. J. Quant. Chem.* **5**, 107 (1971)
- 6b. Flurry, R. L.: *Theoret. chim. Acta (Berl.)* **23**, 1 (1971)
- 6c. Yoshida, Z., Kobayashi, T.: *Theoret. chim. Acta (Berl.)* **23**, 67 (1971)
- 6d. Chesnut, D. B., Moseley, R. W.: *Theoret. chim. Acta (Berl.)* **13**, 230 (1969)
- 6e. Chesnut, D. B., Wormer, P. E. S.: *Theoret. chim. Acta (Berl.)* **20**, 250 (1971)
7. Arnaud, R., Bonnier, J. M.: *J. Chim. Phys.* **66**, 954 (1969)
8. Bonnier, J. M., Arnaud, R.: *J. Chim. Phys.* **68**, 1519 (1971)
9. Diner, S., Malrieu, J. P., Claverie, P.: *Theoret. chim. Acta (Berl.)* **13**, 1 (1969)
10. Malrieu, J. P., Claverie, P., Diner, S.: *Theoret. chim. Acta (Berl.)* **13**, 18 (1969)
11. Diner, S., Malrieu, J. P., Jordan, F., Gilbert, M.: *Theoret. chim. Acta (Berl.)* **15**, 100 (1969)
12. Hoffmann, R.: *J. Chem. Phys.* **39**, 1397 (1963); **40**, 2475 (1964)
13. Pople, J. A., Beveridge, D. L.: *Approximate molecular orbital theory* New York: McGraw-Hill 1970
14. Allen, L. C., Russell, J. D.: *J. Chem. Phys.* **46**, 1029 (1967)
15. Dobosh, P. A.: *Quantum chemistry program exchange*. University of Indiana. Bloomington, Indiana
16. Perahia, D., Pullman, A.: *Chem. Phys. Letters* **19**, 73 (1973)
17. Malrieu, J. P.: *Communication personnelle*.
18. Schwartz, F.: *J. Chem. Phys.* **51**, 4182 (1969)
19. Hillier, I. H., Saunders, V. R., Wyatt, J. P.: *Trans. Faraday Soc.* **66**, 2665 (1970)
20. Bekoe, D. A., Trueblood, K. N.: *Z. Krist.* **113**, 1 (1960)
21. Bonnier, J. M., Arnaud, R.: *C.R. Acad. Sci. Paris, Ser. C* **270**, 885 (1970)

22. Bodor, N., Farkas, M., Trinajstić, N.: *Croat. Chem. Acta* **43**, 107 (1971)
23. Galasso, V., Trinajstić, N.: *Tetrahedron* **28**, 2799 (1972)
24. Wold, S.: *Acta Chem. Scand.* **20**, 2377 (1966)
25. Murthy, A. S. N., Davis, R. E., Rao, C. N. R.: *Theoret. chim. Acta (Berl.)* **13**, 81 (1969)
26. Schuster, P.: *Int. J. Quant. Chem.* **3**, 851 (1969)
27. Gropen, O., Seip, H. M.: *Chem. Phys. Letters* **11**, 445 (1971)
28. Baud, D.: Thèse 3ème cycle, Grenoble 1972
29. Kroll, M.: *J. Am. Chem. Soc.* **90**, 1097 (1968)
30. Briegleb, G., Czekalla, J., Reuss, G.: *Z. Physik. Chem. (Neue Folge)* **30**, 334 (1961)

Dr. M. Gelus  
Université Scientifique et Médicale de Grenoble  
Laboratoire de Chimie Générale  
B.P. n° 53 – Centre de Tri  
F-38041 Grenoble-Cedex, France

## Electronic Structure of Unstable Intermediates

### III. The Electronic Structure of OCC

Colin Thomson and Brian J. Wishart

Department of Chemistry, University of St. Andrews, St. Andrews, Scotland

Received June 22, 1973

The geometry of the species OCC has been investigated within the restricted Hartree-Fock LCAO-MO-SCF approximation. Several one electron properties have been calculated at the calculated minimum energy configuration of  $R(O-C) = 2.121$  bohr,  $R(C-C) = 2.58$  bohr.

**Key words:** Unstable intermediates - Calculated geometry

#### Introduction

This paper reports *ab initio* calculations on the triplet, OCC, and represents another in the series of papers from this Department, dealing with the electronic structure of unstable intermediates. The species has been observed by IR spectroscopy in an argon matrix by Milligan and Jacox [1]. They observed this species 1) on photolysis of matrix-isolated cyanogen azide in the presence of carbon monoxide (at a wavelength less than  $2800 \text{ \AA}$ ) and 2) by vacuum UV photolysis of matrix-isolated carbon suboxide. Features of the IR spectra obtained from these systems were interpreted by Milligan and Jacox as arising from a linear species, OCC.

Our theoretical examination of the system was prompted by their observations, together with the existence of a corpus of work postulating the presence of OCC as an intermediate in a variety of chemical processes, including: radiolysis of carbon dioxide [2]; photolysis of carbon suboxide [3]; thermal decomposition of carbon suboxide [4]; UV photolysis of carbon monoxide [5]; and, pulse radiolysis of carbon monoxide [6].

In the work by Bayes [3], there is reported results of a semi-empirical calculation on OCC, with bond lengths taken from those experimentally determined for carbon suboxide ( $R(C-O) = 2.1921$  bohr;  $R(C-C) = 2.4170$  bohr). These bond lengths were also used by Olsen and Burnelle [7] in another semi-empirical calculation. It is suggested by Bayes that the ground state of OCC is  $^3\Sigma^-$ , with low-lying  $^1\Sigma^+$  and  $^1\Delta$  excited states. Baker et al. [8] showed that their results of photolysis of carbon suboxide in presence of alkenes are consistent with the existence of the two multiplicities (triplet at lower, singlet at higher energies of photolysing radiation).

With the only previous theoretical investigations of this species being semi-empirical, and without geometry optimisation, we have carried out calculations,

determining an optimum equilibrium geometry for the ground state, using *ab initio* LCAO-MO-SCF wavefunctions of similar quality to those used in previous work in this series [9, 10].

### Method of Calculation

Calculations were performed on an IBM 360/195, using the programme AL-CHEMY [9]. This programme computes wavefunctions in the Restricted Hartree-Fock approximation, as formulated for open-shell species by Roothaan [11], for linear molecules, over a basis set of Slater-Type Orbitals (STO's). Included in the programme package are routines for the usual Mulliken population analysis [12], and for computing the expectation values of various one-electron operators. In performing the calculations, the total energy was minimised with respect to each bond length, as described in a previous paper [9].

### Basis Sets

Three basis sets were used in the calculations, at successively greater levels of quality. The initial calculations were carried out with a double-zeta basis set, with exponents taken from the tables of Clementi [13]. This basis set was used to obtain a good first approximation to the bond lengths, after which a further set of calculations were performed using this basis set, augmented by polarization functions of  $3d\sigma$ ,  $3d\pi$ ,  $4f\sigma$ ,  $4f\pi$  symmetry on each atom. This basis we refer to as a  $(DZ + P)$  basis. The polarization function exponents were taken from calculations on  $\text{CO}_2$  with a similar basis set by McLean and Yoshimine [14]. The third basis set was the Bagus-Gilbert "best-atom" set - given by McLean and Yoshimine [14], with the polarization functions taken from their calculation on  $\text{OCN}^-$ . Energy results are reported in hartrees, distances in bohrs.

### Results

The geometries investigated, together with the total energies and the virial coefficients are given in Table 1, for each basis in turn.

Calculations with the double zeta basis gave an optimum geometry of:  $R(\text{O}-\text{C}) = 2.205$  bohr;  $R(\text{C}-\text{C}) = 2.645$  bohr.

For the set of calculations using the  $BA + P$  basis set, the optimum geometry was taken as  $R(\text{O}-\text{C}) = 2.121$  bohr;  $R(\text{C}-\text{C}) = 2.58$  bohr (from total energy vs. bond lengths curves). The total energy was calculated at this geometry, together with expectation values of various one-electron operators. These expectation values are listed in Table 2. At this geometry also, SCF calculations were performed on the anion and cation of the species. Comparison of the orbital and total energies, together with the calculated vertical ionisation potentials are given in Table 3. Summaries of part of the Mulliken population analyses for the three species,  $\text{OCC}$ ,  $\text{OCC}^+$ ,  $\text{OCC}^-$ , are given as Table 4. Values, derived from the expectation values, for observable properties, are given in Table 5.



Table 1. Energy results

Bond O-C	Lengths C-C	Total energy	Virial
Calculations with the (DZ) basis			
2.2	2.63	-150.36982	-2.00027
2.2	2.65	-150.36985	-2.00040
2.2	2.60	-150.36955	-2.00007
2.2	2.67	-150.36976	-2.00051
2.15	2.65	-150.36812	-1.99953
2.25	2.65	-150.36893	-2.00116
2.22	2.65	-150.36977	-2.00071
2.205	2.645	-150.36987	-2.00045
Calculations with the (DZ + P) basis			
2.2	2.64	-150.50279	-2.00205
2.1	2.64	-150.50579	-2.00016
2.1	2.65	-150.50555	-2.00023
2.1	2.645	-150.50568	-2.00019
Calculations with the BA + P basis			
2.10	2.645	-150.51518	-2.00007
2.11	2.645	-150.51538	-2.00028
2.12	2.645	-150.51544	-2.00048
2.14	2.645	-150.51518	-2.00088
2.126	2.645	-150.51541	-2.00060
2.122	2.645	-150.51543	-2.00052
2.122	2.65	-150.51531	-2.00056
2.122	2.55	-150.51615	-1.99975
2.122	2.5	-150.51489	-1.99926
2.122	2.4	-150.50778	-1.99808
2.122	2.3	-150.49262	-1.99601
2.121	2.58	-150.51632	-2.00000

Table 2. Expectation values of quoted operators for OCC

Operator	Value at centre:		
	O	C	C
1/r	26.3283	20.6671	18.7191
z	41.4454	-0.974598	-52.5746
z <sup>2</sup>	178.9141	93.0755	231.2325
r <sup>2</sup>	202.6915	116.8530	255.0099
q <sup>2</sup>	23.7774	23.7774	23.7774
z/r <sup>3</sup>	1.61589	-0.87573	-1.26289
(3z <sup>2</sup> - r <sup>2</sup> )/r <sup>5</sup>	0.72398	1.58626	0.06949



## Relatio

# Gas Phase Acidities of CH Bonds Adjacent to Oxygen and to Sulphur

Saul Wolfe

Department of Chemistry, Queen's University, Kingston, Ontario, Canada

Luis M. Tel

Departamento de Quimica Fisica, Facultad de Ciencias, Universidad de Valladolid, Spain

Imre G. Csizmadia

Department of Chemistry, University of Toronto, Toronto, Ontario, Canada

Received June 27, 1973

It has been shown, by double zeta quality *ab initio* MO calculations, that, in the gas phase, a CH bond adjacent to sulphur is more acidic than a CH bond adjacent to oxygen. This trend is in agreement with experimental observations in solution.

**Key words:** Gas phase acidity – Proton affinity – Carbanions

A hydrogen-containing molecule (AH) may undergo either protonation or deprotonation:



and the relative facilities of the two processes are measured, respectively, by the gas phase basicity and the gas phase acidity of the molecule. Since proton affinity ( $A_H^+$ ) is defined, either in the thermodynamic convention:

$$\begin{aligned} A_H^+(\text{MH}) &= \Delta H_f^0(\text{MH}_2^+) - \{\Delta H_f^0(\text{MH}) + \Delta H_f^0(\text{H}^+)\} \\ A_H^+(\text{M}^-) &= \Delta H_f^0(\text{MH}) - \{\Delta H_f^0(\text{M}^-) + \Delta H_f^0(\text{H}^+)\}, \end{aligned}$$

or in the quantum mechanical convention:

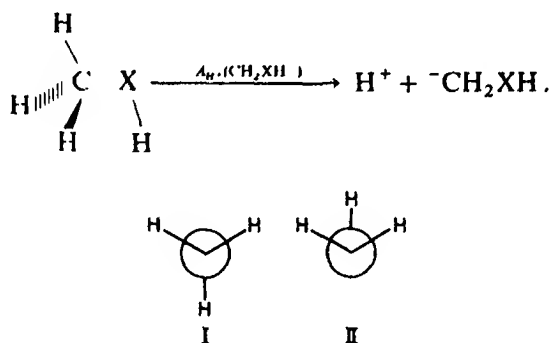
$$\begin{aligned} A_H^+(\text{MH}) &= E(\text{MH}_2^+) - E(\text{MH}) \\ A_H^+(\text{M}^-) &= E(\text{MH}) - E(\text{M}^-), \end{aligned}$$

as a difference between protonated and deprotonated states, the gas phase basicity and acidity of the molecule may be defined as:

$$\begin{aligned} \text{gas phase basicity of MH} &= A_H^+(\text{MH}), \\ \text{gas phase acidity of MH} &= -A_H^+(\text{M}^-). \end{aligned}$$

Introduction of a heteroatom into a hydrocarbon molecule leads to a change in the acidities of the remaining C–H bonds [1]. Because the quantitative effect of each heteroatom is different, a quantitative understanding of the effects of heteroatom substitution upon carbanion proton affinities is a necessary prerequisite to the understanding of experimental acidity differences. For example, it is an experimental fact [2] that the kinetic acidity of a proton at a carbon adjacent to sulphur is several orders of magnitude greater than that of a proton at carbon adjacent to oxygen. This has been attributed to  $d_{\pi}$ – $p_{\pi}$  stabilization of the anion in the former case [2, 3]. However, it has not been necessary to postulate  $d$ -orbital conjugation to account for the *stereochemical* properties of sulphur-stabilized carbanions [4], and it seemed to us, therefore, that some, as yet unrecognized, non-conjugative explanation might also account for the *chemical* difference between the two systems.

The approach that is being taken involves double zeta quality computations of the energies of  $\text{HOCH}_3$ ,  $\text{HOCH}_2^-$ ,  $\text{HSCH}_3$ , and  $\text{HSCH}_2^-$ , with full optimization of the geometries, and with  $s$ ,  $sp$ , and  $spd$  basis functions on the heteroatoms. In the first part of the work, geometry optimization of  $\text{HOCH}_3$  [5] and  $\text{HOCH}_2^-$  [6] was achieved for  $sp$  bases. This permits evaluation of the following gas phase acidity, for  $\text{X}=\text{O}$ :



In the present work, the gas phase acidity for  $\text{X}=\text{S}$  has been computed with an  $spd$  basis. Bond angles about the carbon atom in the carbanion have been assumed to be tetrahedral. Although full geometry optimization has not yet been performed, the trends in the gas phase acidities are already clear.

Table 1. Conformational energies of  $\text{HOCH}_2^-$ ,  $\text{HOCH}_3$ ,  $\text{HSCH}_2^-$ , and  $\text{HSCH}_3$

Rotational angle <sup>a</sup>	0°	60°	120°	180°
$\text{HOCH}_2^-$	–114.30505	–114.29316	–114.28616	–114.29469
$\text{HOCH}_3$	–115.00875	–114.01105		
$\text{HSCH}_2^-$	–437.01398	–437.00044	–437.00131	–437.01555
$\text{HSCH}_3$	–437.68839	–437.69026		

<sup>a</sup> For definition of these angles, see Fig. 1.

The data are summarized in Table 1 and illustrated in Fig. 1. Two minima are found for each carbanion. These correspond to the *Y(I)* and *W(II)* conformations. In  $\text{HOCH}_2^-$  the *Y* conformation represents the lowest minimum, but in  $\text{HSCH}_2^-$  the *W* conformation has the lower energy (Fig. 1). The gas phase acidities, also shown in Fig. 1, are:

$$-A_{\text{H}}^+(\text{HOCH}_2^-) = 0.70370 \text{ hartree} = 441.7 \text{ kcal/mole},$$

$$-A_{\text{H}}^+(\text{HSCH}_2^-) = 0.67284 \text{ hartree} = 422.3 \text{ kcal/mole}.$$

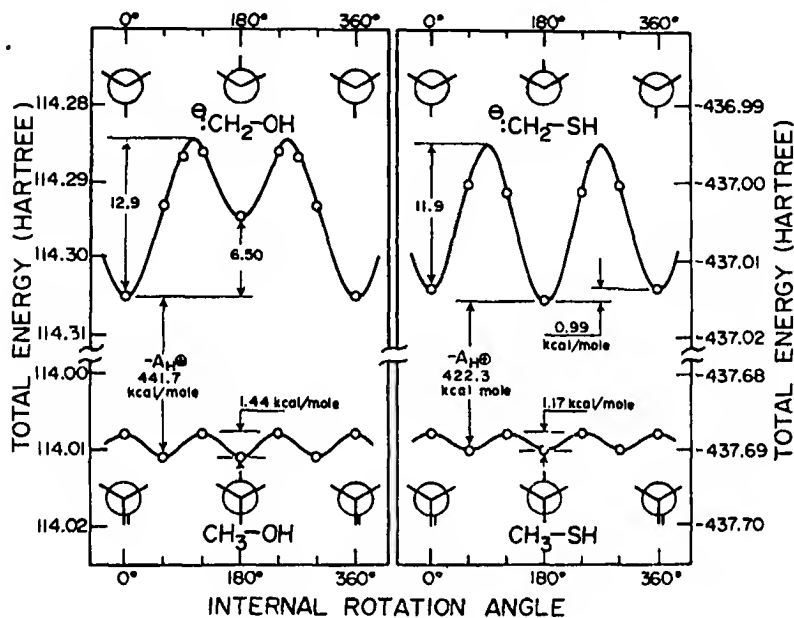


Fig. 1

These values indicate that a CH bond adjacent to sulphur is more acidic than a CH bond adjacent to oxygen. The final account of this work will contain the full geometry optimization of  $\text{HSCH}_2^-$  and the comparative study of the gas phase acidities as a function of basis set.

**Acknowledgements.** We thank the National Research Council of Canada for financial support of this work, and Professor A. Mangini, Università di Bologna, for his kind hospitality during the preparation of this manuscript.

### References

1. Cram, D. J.: *Fundamentals of carbanion chemistry*. New York: Academic Press 1965
2. Oae, S., Tagaki, W., Ohno, A.: *Tetrahedron* **20**, 417 (1964)
3. Coffen, D. L.: *Rec. Chem. Progr.* **30**, 275 (1969); – Mitchell, K. A. R.: *Chem. Rev.* **69**, 157 (1969)
4. Wolfe, S.: *Accounts Chem. Res.* **5**, 102 (1972), and references cited in this article
5. Tel, L. M., Wolfe, S., Csizmadia, I. G.: *J. Chem. Phys.* **59**, in press (1973)
6. Wolfe, S., Tel, L. M., Csizmadia, I. G.: *Can. J. Chem.* **51**, 2423 (1973)

Prof. Dr. I. G. Csizmadia  
University of Toronto  
Dept. of Chemistry  
Lash Miller Laboratories  
80 St. George St.  
Toronto, Canada

